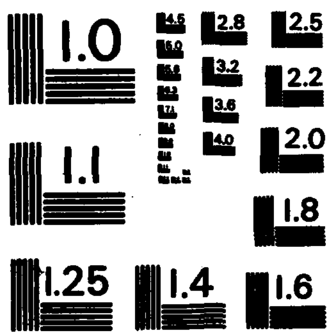
AD-A122 064 SIMULATOR DESIGN FEATURES FOR CARRIER LANDING I PERFORMANCE EXPERIMENTS(U) CANYON RESEARCH GROUP INC WESTLAKE VILLAGE CA D P WESTRA ET AL. DEC 81 1/1

UNCLASSIFIED CRG-TR-81-016 N61339-78-C-0060 F/G 5/9 NL

END

FILMED

DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

(12)

ADA122064

Technical Report NAVTRAEQUIPCEN 78-C-0060-7

SIMULATOR DESIGN FEATURES FOR CARRIER LANDING: I. PERFORMANCE EXPERIMENTS

INTERIM REPORT 1 MAY 1980 - 31 AUGUST 1981

September 1982

DoD Distribution Statement

Approved for public release; Distribution unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NAVTRAEQUIPCEN 78-C-0060-7	2. GOVT ACCESSION NO.: AD-A122064
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): SIMULATOR DESIGN FEATURES FOR CARRIER LANDINGS: I. PERFORMANCE EXPERIMENTS	5. TYPE OF REPORT & PERIOD COVERED: 1 May 1980 - 31 August 1981
	6. PERFORMING ORG. REPORT NUMBER: TR-81-016
7. AUTHOR(s): D. P. Westra; C. W. Simon; S. C. Collyer; W. S. Chambers	8. CONTRACT OR GRANT NUMBER(s): N61339-78-C-0060
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Canyon Research Group, Inc. 741 Lakefield Road, Suite B Westlake Village, CA 91361	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 4781-6P1A
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Training Equipment Center Orlando, Florida 32813	12. REPORT DATE: December 1981
	13. NUMBER OF PAGES: 62
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Visual Technology Research Simulator (VTRS); Multifactor experiments; Holistic experimention; flight simulation; carrier landing research; visual simulation; motion simulation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The effects of twelve factors on carrier landing performance in the Visual Technology Research Simulator (VTRS) at the Naval Training Equipment Center (NTEC), Orlando, Florida, were investigated in a series of three experiments. Subjects for the experiment were experienced naval aviators. The purpose was to determine and rank order the sizes of effects, identify factors having no effect, and to obtain information for making decisions about future transfer-of-training studies.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

cont'd

20. ABSTRACT (Cont'd)

In the first experiment, the task was a straight-in approach and landing. Seven visual display factors (Fresnel Lens Optical Landing System type, ship detail, field of view, visual lags, seascape, brightness, TV line rate), two non-visual factors (motion and engine lags), one environmental factor (turbulence) and subjects as a factor were studied. In the second experiment, a circling approach to landing was employed as the task and included as factors ship detail, visual lags, seascape, brightness, motion, and turbulence. In the third experiment, a straight-in approach was employed. Two simulation factors, G-seat and ship type, were studied along with turbulence.

Results generally showed small to null effects for equipment factors although several had statistically reliable effects. As the display and simulator factors were manipulated over a wide range of interest representing expensive vs. inexpensive simulation options, the implication is that simulation for carrier landing skill maintenance and transition training for experienced pilots does not require the highest levels of fidelity for these features. Simulator requirements for training at the undergraduate level are currently being examined.

ACKNOWLEDGEMENTS

The conduct of the experiments reported here required the dedicated efforts and support of a number of personnel. The entire technical staff at VTRS deserves special mention for providing system capability to support the rapid factor level switching required by the multifactor experimental designs. Brian Nelson performed most of the data analysis and Lee Wooldridge of Vreuls Research Coproration, Inc. provided consulting and assistance for multivariate data analyses.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF TABLES (cont.)

APPENDIX A

SECTION I

INTRODUCTION

The Visual Technology Research Simulator (VTRS) at the Naval Training Equipment Center (NTEC), Orlando, Florida, is designed for research on flight simulator requirements for training and skill maintenance. The VTRS consists of a fully instrumented Navy T-2C jet trainer cockpit, a six-degrees-of-freedom synergistic motion platform and a wide angle visual system that can project both computer-generated (CIG) and model-board images onto a spherical screen. The visual system is capable of displaying images via target and background projectors subtending 50° above and 30° below the pilot's eye level and can display 160° of horizontal field (Collyer and Chambers, 1978).

The current phase of effort at VTRS involves research to define simulator requirements for the carrier landing task. Because of the large cost implications there is a need to investigate a large number of visual and other simulator features. A research program was planned around the holistic experimental philosophy and paradigm proposed by Simon (1973; 1977). Fundamental to this approach is the importance of studying as many factors of interest as possible within a single experiment.

The experiments reported here investigated the effects of eleven simulator variables and their critical interactions on the performance of experienced pilots making straight-in and circular approaches for carrier landings. This research is the first phase in a program that will include "quasi-transfer" studies, in which the simulator is both the training and criterion vehicle, as well as transfer studies involving actual flight tests. The information obtained in this first set of studies is directly relevant to the design of simulators for skill maintenance and for transition training. These two types of training are considered to be substantially more expensive than undergraduate pilot training (Orlansky and String, 1977).

This first series of three experiments, however, served several additional purposes. It provided a vehicle for exhaustively testing and debugging both the hardware and software of the new simulator as a research tool and as a training device. It provided a basis for developing better performance measures and an improved methodology for conducting equipment design research. The information obtained from the skill maintenance studies is useful in planning experiments at later stages in the program. The findings also can be compared with those obtained later in transfer investigations with less experienced pilots to begin to establish general principles for relating performance to transfer.

Three experiments were performed. The two primary experiments included most of the factors of interest and involved straight-in and circular approaches, respectively. The third was an adjunct study involving two factors that could not be readily manuipulated within the other experiments.

SECTION II

EXPERIMENTAL PLAN

General characteristics of the three experiments are described below. Speci^ic details of each experiment are provided in Sections III-V.

MISSION

Pilots were assigned the mission of making a daytime carrier approach and landing a T-2C jet on the deck of the aircraft carrier Forrestal which was moving at 20 knots with a zero effective crosswind over the landing deck and 25 knots relative wind down the deck. The complete circling approach and landing is depicted in Figure 1. The landing was made without the aid of a landing signal officer (LSO) and no wave offs were given.

The daytime mission was emphasized in part because of an interest in the issues surrounding the need for a wide-angle display. Night carrier landings involve straight-in approaches, and several night trainers now in use have relatively narrow fields of view (approximately ±25 degrees horizontally). The daytime mission would, on the other hand, require a much larger field of view if it were considered essential for the pilot to see the ship continuously during the circling approach. Additionally, daytime training raises more pressing questions of required scene content. Since nighttime operations are much easier and less costly to simulate, the effects of reduced scene fidelity in night displays are seldom at issue.

TASKS. The carrier-landing task was separated into what are considered to be behaviorally distinct components: "turn" and "final approach", and the data were treated separately.

Straight-In Task. The experimental trial was initialized with the aircraft at 7200 feet from the ramp on the glideslope, on centerline with the landing deck and in the approach configuration (full flaps, speed brake out, hook and wheels down and 15 units angle of attack). Fuel was fixed at 3200 pounds to give a gross aircraft weight of 10,000 pounds. A trial terminated with either a successful wire arrestment, a bolter (i.e. touchdown past the wire arrestment area), or no touchdown flown either to crash or 1000 feet past the carrier. This task was used in Experiments I and III.

Circling Task. The aircraft was initialized abeam the LSO platform on the downwind leg (see Figure 1) at 6200 feet from the ship and at 600 feet of altitude in the approach configuration. A trial consisted of the final turn, final approach and attempted landing. With circling approaches, turn performance could be examined separately and the effect of turn performance on final approach and landing could be determined. This task was used in Experiment II.

SEGMENTS. Performance summary measurements were made along the flight path for specified segments conforming roughly to traditional Navy designation of "start", "in the middle", "in close", and "at the ramp". Major segments subjected to separate analyses are given below.

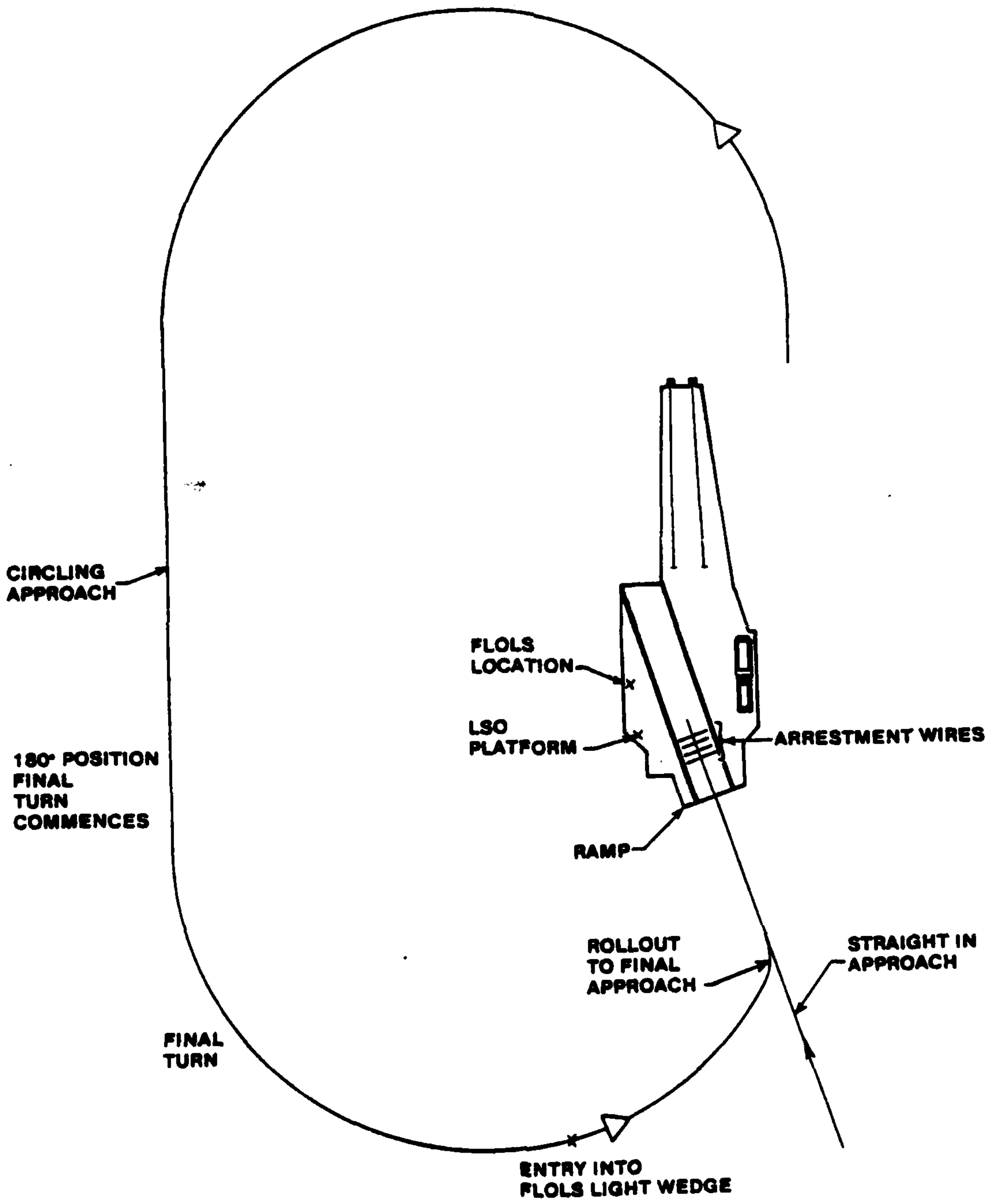


Figure 1. Overhead View of Typical Daytime
Circling Carrier Landing Pattern and Night Straight-in Approaches.

Circling

a. Entry segment (after initialization): Between the Fresnel Lens Optical Landing System (FLOLS) space entry point and rollout to final approach. The FLOLS is described in detail later in this report. The FLOLS spacr entry point is the location of the aircraft in the turn when it first enters the FLOLS light wedge (see Figure 1) and where the pilot can first see the FLOLS lens. The point of entry will vary somewhat depending on aircraft altitude but will generally be just past the 90° point in the final turn. The rollout point was defined as the point at which the aircraft heading had crossed the ship's heading and aircraft roll altidude was within ±4° of level.

b. In-close segment: 2000' to 500' from the ramp.

c. At-the-ramp segment: 1000' to the ramp.

Straight-In

a. Entry segment: 6000' to 4500' from the ramp.

b. Middle segment: 4500' to 2000' from the ramp.

c. In-close segment: 2000' to 500' from the ramp.

d. At the ramp: 1000' to the ramp.

In addition, measurements were taken at a number of designated "capture" points along the flight path including the point of touchdown. Specific performance measures are discussed later in this section.

FACTORS AND LEVELS

A successful carrier approach and landing involves the use of a family of visual cues external to the cockpit. The principal cues come from a visual landing aid called the Fresnel Lens Optical Landing System (FLOLS) for vertical glideslope control, the carrier deck runway, centerline and dropline markings for lineup control, and the sky, horizon, and seascape for general aircraft attitude control. Other cues necessary to the operation of the aircraft, possibly including motion, are also involved in carrier approach and landings.

A large number of factors potentially affecting these cues were tentatively selected as candidates for study. These were pared down by a panel of engineers and psychologists into the set of twelve factors shown in Table 1. This final list of factors represented a number of issues relevant to the design of carrier-landing trainers. They were considered to be the most important in terms of cost implications and potential effects. These factors could generally be categorized as display and simulator hardware fidelity variables. Eight of these factors were visual display system parameters which directly affected the quality of visual cues. Training aids or augmented simulation features were not considered for this performance study as they could not be meaningfully studied apart from their effect on learning and transfer of training.

TABLE 1. EXPERIMENTAL FACTORS AND LEVELS FOR EXPERIMENTS I, II, AND III

MISSION: Carrier Landing Task

PILOT EXPERIENCE: High

TASK:
Experiment I. Straight-In
Experiment II. Circling
Experiment III. Straight-In

FACTOR	LEVEL SETTINGS "low"	LEVEL SETTINGS "high"	EXPERIMENT I	EXPERIMENT II	EXPERIMENT III
FLOLS	•TV/CIG	Optical/Model	X		
Field of View	-27°: + 9° Vertical ±24° Horizontal	•-30°: + 50° Vertical ±80° Horizontal	X		
TV Line Rate	525	•1025	X		
Engine Lags	7.5 Hz Update	•30 Hz Update	X		
Ship Detail	Night point-light	•Day Solid Surface	X	X	
Visual Lags	200 msec total delay	•100 msec total delay	X	X	
Seascape	Gray Homgeneous Background	•Wave Pattern	X	X	
Brightness	Ship: 0.40 fL Sea: 0.04 fL Sky: 0.02 fL	•2.90 fL 0.50 fL 0.16 fL	X	X	
Motion	•Fixed Base	Six-Degrees-of-Freedom	X	X	
Turbulence	Close to Maximum Flayble	No Turbulence	X	X	X
Ship Type	•CIG	Camera/Model Board			X
G-Seat	•Off	30 pneumatic bellows			X

Notations: X in an Experiment column indicates that the factor aligned with it was actually varied; i.e., both levels were studied.

No X indicates that the factor was held constant in the particular experiment at the level setting preceded by the large black dot (•).

All factors were not included in all three experiments. The ones that were varied in each experiment are designated by X's in Table 1. Experiment I was viewed as the primary study with Experiments II and III providing supplemental data on task segments or factors that would have been inappropriate or disruptive to have included in Experiment I, but which were considered important to the program. Only the factors likely to affect turn performance and the subsequent carrier landing were included in Experiment II. The variable called Ship Type was not included in the first two experiments because it was feared that unreliability of the carrier-model system plus the time required to shift from it to the CIG system might compromise the other experimental data. The effects of the two ship image types were therefore studied in the third experiment.

High and low factor settings were chosen in order to bracket the reasonable range of interest. For the equipment factors, the high levels were generally set at the highest state-of-the-art engineering levels attainable while the low levels were chosen to be the most degraded form of the factor likely to be employed operationally.

FLOLS. The FLOLS provides glideslope displacement information to the pilot during an approach. Physically located forward of the LSO platform and on the port side of the landing deck (see Figure 1), the FLOLS consists of five Fresnel lenses vertically arranged between two horizontal light arrays known as datum bars. The array of Fresnel lenses provides an image which appears to the pilot as a single sphere of light known as the "meatball". This meatball is visible to the pilot within a wedge of space .75° above and below the projected glideslope of 3.5° and ±25° horizontally from the center of the wedge projected parallel to the landing deck. The pilot judges his angular glideslope deviation from the distance the meatball appears to be above or below the datum bars. A meatball that appears centered vertically between the datum bars indicates to the pilot that he is on the proper glidepath. Figure 2 gives a view of the FLOLS and its projection aft of the carrier and the perceived relationship of the meatball to the datum bars. Golovcsenko (1976) and Kaul, Collyer and Lintern (1980) provide more detail on FLOLS geometry.

The FLOLS high level was an optical projection (Singer-Link, 1977) of a scale model of the FLOLS providing essentially real-world duplication of the 24 foot wide array of lights except for a constant magnification of 1.5X. This high resolution direct optical projection was specially built for this test to represent maximum realism. It is not used in current trainers and would be a high cost additional subsystem. The FLOLS low level was a TV projection of a CIG (General Electric, 1979) FLOLS data base defined by 96 edges. The two datum bars which normally consist of six lights each were represented by two solid bars. At long range the FLOLS was magnified three times normal size, gradually shrinking to 1.5X at the ramp. The large magnification was required to compensate for limited TV resolution so that the pilots could discriminate meatball position for glideslope guidance at a range similar to that in the real world. This FLOLS magnification technique is used regularly on Navy carrier-landing trainers. For example, the TA-4J trainer (device 2B35) uses magnifications as large as 7X. This technique involves no added cost to CIG carrier-landing visual equipment.

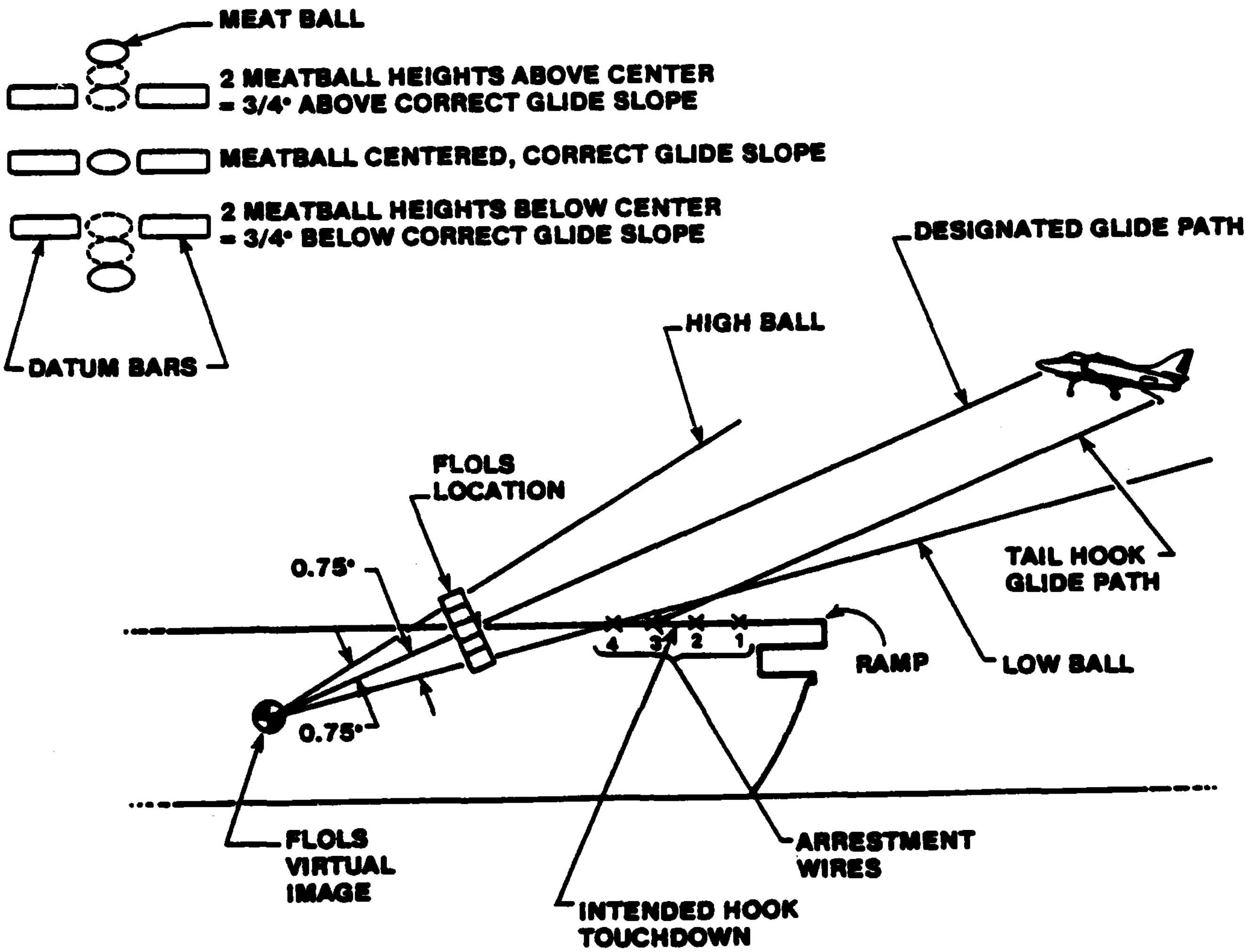


Figure 2. Carrier Approach Geometry Depicting FLOLS Projection of Glideslope Deviation Information. Adapted from Golovcsenko (1976).

SHIP DETAIL. The ship detail high level was a daytime solid model CIG (General Electric, 1979) carrier whose surfaces were defined by 985 edges. It contained all landing deck markings and was scaled to normal size. This detail level was approximately representative of that available from daytime CIG systems costing several million dollars such as the 2B35 trainer, although displayed at higher resolution than available in the 2B35. The ship detail low level was an image of a night point-light CIG carrier consisting of 137 lights. It contained all deck outline, runway, centerline, and drop lights, with the runway surface becoming visible during the final 1600 feet of the approach. Its image is representative of the ship in a night CIG system costing less than a million dollars and used on several Navy Night Carrier-Landing Trainers (NCLT's).

FIELD OF VIEW. The field of view high level was a 160° horizontal by 80° vertical wide-angle display (Singer-Link, 1977) which is costly and is representative of that currently available for carrier-landing training only on multitask trainers such as the 2B35 and the F-14 Wide-Angle Visual System (WAVS). For circling approaches the horizontal field of view was offset 40° to the left to give -120° to +40° of viewing area. The field of view low level was a ±24° horizontal by -27° to 9° vertical narrow-angle display which is representative of the lower cost NCLTs used for F-4, F-14, A-6, A-7, and S-3 transition training.

VISUAL LAG. The visual lag (Browder and Butrimas, 1981) high level of 100 msec from stick input to the completion of the first field of video output was representative of a 30 Hz update computer simulation and a 60 Hz CIG update. This response time is faster than that available on current Navy trainers and represents increased computer capacity and cost. The visual lag low level of 200 msec is representative of a 15 Hz update computer simulation and 30 Hz CIG update. This response time is representative of the lowest level normally considered in acquisitions of simulators with visual systems.

SEASCAPE. The seascape high level was a water texture pattern generated by a flying spot scanner (Singer-Link, 1977) readout of a photographic film plate of a seascape wave pattern. It provided translation and ground growth as well as attitude (i.e., roll, pitch, yaw) information which is considered to be potentially useful as cues for improved aircraft control and visual realism. The seascape low level was the absence of the texture pattern and a uniform grey below the horizon, providing roll and pitch information only. This represents a reduced cost by eliminating the seascape image generator.

BRIGHTNESS. The brightness (Owen, 1980) high level of 2.9 foot-Lamberts (fL) for the target is representative of narrow angle visual systems used on trainers. The brightness low level of 0.4 fL is representative of wide-angle dome displays such as the F-14 WAVS. To increase the wide-angle display brightness to the higher level, if found advantageous for carrier-landing training, would be very expensive and would require considerable development time.

LINE RATE. The TV line rate high level of 1025 is the maximum available in standard TV equipment. The TV line rate lower level of 525 provides only half the resolution but is representative of low-cost commercial equipment. The

factor as used in this experiment relates only to narrow-angle target projectors which use zoom magnification to provide a high resolution image of the carrier but not the background scene.

SHIP TYPE. The ship image type high level was a three dimensional scale model (Singer-Link, 1977) of the Forrestal Carrier CVA-59 scaled at 370 to 1. It contained all deck markings including tire skid marks on the runway touchdown area as well as ship details such as gun mounts, antennas and railings. It was viewed by a TV camera with a four-to-one zoom lens mounted on a computer-controlled gantry and optical probe. The ship type low level was the daytime CIG carrier CVA-59 model described under the ship detail paragraph. The 3-D model ship type is usually considered to be of high fidelity and is used in several engineering and research simulators. While this technique has been built in several past prototype trainers it has not and is not being used in a major carrier-landing trainer.

MOTION. A six-degrees-of-freedom 48 inch synergistic motion platform (Browder and Butrimas, 1981) was fully operational for the high level, and was stationary for the low level of this factor. This platform is similar to those on the Navy's 27 T-2C Instrument Trainers (device 2F101) used in Undergraduate Pilot Training (UPT) except that VTRS computation rates are higher for reduced cuing time lag. While it is representative of many older platforms on existing trainers, it does not have the low noise and improved response of new platforms.

ENGINE LAG. Engine computation time lags were set at 30 Hz and 7.5 Hz for high and low levels, respectively. This factor was included because the rather large engine computations are typically done at low rates such as 7.5 Hz to minimize computer size and cost. Preliminary testing had generated positive reactions from pilots who indicated that the 30 Hz engine updates caused the simulator response to feel more realistic. This was considered important since Navy carrier-landing approach procedures for jet aircraft typically require frequent throttle changes for glideslope control.

G-SEAT. A G-seat (Browder and Butrimas, 1981) having 30 pneumatic bellows was operational for the high level and stationary for the low level. The seat was operated with the normal software used to simulate high sustained g-cues and did not contain any special gain setting for a landing task. The seat design is similar to that used in the ASPT, SAAC and several F-4 and F-14 trainers. Its software computations were performed at a 30 Hz rate.

TURBULENCE. Turbulence was included in the experiment to allow examination of factor effects under two difficulty levels. The high level was set at no environmental turbulence. The low level was set at turbulence in the form of wind acting on the longitudinal, lateral and vertical aircraft axes. These "winds" were generated by the summation of sine waves varying in frequencies and amplitudes. They were strong enough to create a degree of turbulence judged by Navy pilots to be near the maximum level under which operations would proceed at sea.

PERFORMANCE MEASURES

The complex nature of the carrier-landing task requires that a variety of measurements be used to describe and evaluate performance under the range of simulation conditions derived from the multiple factors and the differences in tasks and segments of the complete mission. The selection of the performance measurements collected for these experiments was based on a literature search and VTRS preliminary data collection and analysis (Lintern, et al., 1979; Lintern, et al., 1980). A list of all summary measures collected is given in Appendix A.

The initial list was pared down on a rational basis until a working set of variables was obtained. This list was first reduced by examining the table of intercorrelations to find those variables correlating .90 or higher. Of those, the one having the highest multiple correlation with all variables was eliminated. Then all remaining multiple correlations were examined and those having a multiple correlation of .95 or higher with all of the other remaining variables were removed. Subsequent data analysis was performed only on the measures from this reduced set which differed slightly from experiment to experiment.

The specific measures used in all analyses are:

RMS ERROR. Three root-mean-square (RMS) error measurements of aircraft position relative to the optimum final approach flight path for a safe landing were used, i.e.:

- Vertical deviations in degrees from the specified glideslope;
- Horizontal deviations in degrees from the center of the landing deck (lineup); and
- Angular deviations, in units, from the optimum angle of attack.

In addition, RMS error was separated into its independent bias and variable components and examined. These measures were recorded over several segments of the approach and descent to determine whether or not effects differed at different distances from the ship.

PERCENT TIME-ON-TARGET. A percent time-on-target measure was calculated during the last thousand feet to the ramp for each of the above measures. Based on recommendations by Navy Landing Signal Officers (LSOs), tolerance bands defining acceptable performance were set at $\pm .3^\circ$, $\pm 1.5^\circ$, and ± 1.0 units for glideslope, lineup, and angle of attack, respectively. The tolerance for glideslope represents approximately plus or minus one "meatball" of deviation of the FLOLS display.

TOUCHDOWN SCORES. Touchdown scores were made for these five performance measures at the point of touchdown:

- o Wire trapped,
- o Lineup deviation from center of landing deck (Feet),
- o Vertical velocity (Ft/Sec),
- o Aircraft pitch (Degrees), and
- o Aircraft roll (Degrees).

Means and dispersion values were calculated. The dispersion values are standard deviations calculated across all experimental conditions and thus include the effects of other sources of variance. They are presented here for descriptive purposes only.

TOUCHDOWN QUALITY SCORES. Touchdown quality scores were also calculated. These scores and their limits are given in Table 2. They were a transformation of the above list of actual measurements into scores representing ratings of the accuracy and safety of the aircraft position at the point of touchdown. These "touchdown quality scores" were scaled from zero to 100 with zero indicating performance that was at the point of causing damage to the aircraft or ship, was totally outside of acceptable limits, or was an approach that did not result in touchdown on the landing deck. A score of 100 indicated performance within a fairly narrow range representing "perfect" performance. This was defined as performance within ± 1/2 standard deviation (empirically obtained from these experiments and preliminary work) from task defined optimum values for each measure. Scores between the "perfect" and unacceptable limits were linearly interpolated from 100 to zero. Only the wire-trapped quality score was derived differently.

WIRE TRAPPED QUALITY SCORE. The wire-trapped quality score of 100 was assigned to a #3 wire catch, 75 for the farthest point from the ramp resulting in a #4 wire catch, and 58.3 for a hook touchdown 10' aft of the #1 wire. These values are partially based on Brictson's (1973) landing performance score. Scores were linearly interpolated between these points and between the limits of 40' aft of the #1 wire and 80' fore of the # 4 wire which were assigned scores of zero. The asymmetrical assignment of scores about the #3 wire reflects the fact that it is more dangerous to land short (aft of the wires near the ramp) than it is to land an equivalent distance long (beyond the wires). Consideration of the landing area before and beyond the wires in the scoring was necessary since pilots were not waved off from the approach and almost always attempted a landing. The assumption is made that an unacceptable landing in the simulator (outside the limits described above) would have been waved off in the real world. Scores of zero were assigned to all touchdown quality indices when an approach did not result in touchdown on the landing deck.

PERCENT SUCCESSFUL LANDINGS. The percent of successful landings on the carrier deck was determined for each condition. This score takes into consideration the three most important indicators of touchdown quality (wire trapped, lineup, and vertical velocity). It represents an attempt to provide

TABLE 2. SCALING TOUCHDOWN QUALITY MEASURES

Lineup Score:	100 if < ' ± 5' ' 0 if > ' ± 18' '
Roll Score:	100 if < ' ± 1.5° ' 0 if > ' ± 9° '
Pitch Score:	100 if > 6.63° and < 8.63° 0 if < 2.63° or > 12.63°
Vertical Velocity Score:	100 if > 7.05 ft/sec and < 9.45 ft/sec 0 if < 0.83 ft/sec or > 15.67 ft/sec
Wire Trapped Score:	100 if #3 wire trap 0 if > 40' aft of #1 wire or > 80' fore of #4 wire or if no touchdown was made

an "overall" indicator of terminal performance. "Success" is specifically defined here as being a landing in which wire 1, 2, 3, or 4 was caught and touchdown occurred no more than 30' aft of the first wire with lineup within 18 feet of the center-line, and vertical velocity not greater than 15.67 ft/sec. Outside of these limits, under operational conditions, it is assumed that the attempt would have been waved off or would have been a bolter.

OTHER MEASURES. Other performance measurements were also made for this task. Aircraft output measures, i.e., pitch and roll, were measured in terms of both bias and variable error over the segments of flight. Pilot input measures, such as total stick movement for aileron, elevator, throttle, and pedal per segment, were also recorded but proved of little significance and will not be discussed further.

PILOTS

The eight pilots used as subjects in the three experiments were experienced Naval aviators with at least one tour at sea. The average number of military flight hours for these pilots was 2254, ranging from 630 to 4500. The average number of carrier landings (non-simulator) was 346, ranging from a low of 50 to a high of 800. Five of the pilots had their most recent flight experience in the A-7 aircraft, two in the S-3 and one in the A-4. All were stationed at Naval Air Station Cecil Field, Florida, at the time of the research.

SCHEDULING

On the first day, pilots received approximately 32 pre-training trials, half with straight-in approaches and half with circling approaches. Some pilots received up to 10 extra trials if their performance was erratic or if they were still unsatisfied with their performance. This rather extensive pre-training was done to stabilize flight strategies within pilots as much as possible. Pilots were instructed to try to stabilize their performance during pre-training and to perform as consistently as possible during the experiments.

Each of the eight pilots performed all three experiments consecutively: 32 trials on Experiment I on the second day, 32 trials on Experiment II on the third day, and 16 trials on Experiment III on the fourth day. Each pilot performed 16 consecutive trials in a single simulator session. A pilot performed one morning and one afternoon session per day until finished, each session lasting about one hour.

PILOT OPINIONS

At the end of the study, pilots were asked to rate each equipment factor level with regard to its fidelity, adequacy for training, and adequacy for skill retention. Pilot experience data in terms of number of flight hours and actual carrier landings were also obtained.

DATA ANALYSIS

HANDLING OUTLIERS. The performance data from each experiment were edited for deviant trials. A trial was defined as deviant if it had more than three measures greater than four standard deviations from the mean of all data for those measures or had one measure greater than five standard deviations from the mean. This resulted in the deletion of five trials from Experiment I data, two from Experiment II and two from Experiment III. Several of the deleted trials were those in which the pilot reported he had lost control prior to reaching the ramp indicating unacceptable data in terms of the task, i.e., the pilot was doing something other than attempting to land on the carrier. Values for the deleted trials were added back to the data from Experiment I by substituting a value approximating the level of the largest non-deviant observation. Data from Experiments II and III were analyzed without substituting values for the deleted trials.

MULTIFACTOR, UNIVARIATE ANALYSIS. The effect of each source of variance was isolated for each performance measure individually using standard analysis of variance procedures. This included all main effects, including turbulence, two-factor interactions, block effects (probably the result of differences in pilot performance), and a residual. In some cases, only combined rather than individual two-factor interaction effects are reported in the tables. Mean differences between the high and low levels of each main effect were determined along with the proportion of total variance accounted for by each source of variance. Since half of the total number of trials in the experiment are used to obtain each mean, each mean difference is based on two sets of 128 observations in the first two experiments and 64 in the third.

Tests of statistical significance were performed and those effects that exceeded a selected probability level were identifed with asterisks. For Experiment I, the probability value used was .005, for Experiment II, .01, and for Experiment III, .05. Significance levels were selected to compensate for the large number of analyses that were being made, the large number of factors in the experiment, and the large sample size involved. The values were chosen to avoid giving too much weight to results that are not likely to be repeated. Those effects without asterisks are not likely to be critical, neither statistically nor practically, although occasionally effects of marginal signficance are noted. Where pre-selected analyses were done for evaluating limited and specific sources of variance, different probability values for significance were occasionally selected and duly noted.

OTHER INFORMATION. There are certain computations that can be performed to obtain supplementary information with which to interpret the data. The numbers required for these analyses are available in the tables.

The mean performance for high and low levels of any factor can be obtained by taking the grand mean shown at the bottom of each column and to it add (high level) or subtract (low level) half of the mean difference for that factor. The sign of the mean difference must be taken into consideration in this calculation since the mean of the low condition was always subtracted from the mean of the high condition to obtain the mean difference. A negative RMS mean difference implies better performance under the high level of a

factor. On the other hand, a positive percent time on target mean difference indicates better performance with the factor high level.

Confidence limits for the mean differences can be roughly obtained by multiplying the standard error of the mean difference -- "STD ERR DIFF" at the bottom of the table -- by plus or minus two for the 95% level and plus or minus 2.6 for the 99% level. These values can be added to the mean differences of each factor to obtain the low and high limits within which the true mean difference is expected to lie.

F-ratios for a particular effect can be calculated using the percentages in the tables. The numerator of the ratio is the percent variance accounted for by an effect divided by its degrees of freedom and the denominator is the residual percent variance accounted for divided by its degrees of freedom. The interested reader may substitute other probability levels than the ones selected here for these F-ratios.

INTERPRETATION STANDARDS. That an observed difference between two conditions is or is not statistically significant provides little or no information regarding the practical significance of the difference. Some outside, "real world" standards are needed to better evaluate the data. For one thing, it would be desirable to relate approach RMS error scores to the terminal criterion, i.e., making a safe carrier landing. In spite of the popular use of the RMS error measure, interpreting it is difficult. For example, there were correlations of about 0.50 between adjacent non-overlapping segments for RMS error measures of the same variables. There were only low correlations (approximately 0.20) between non-adjacent final approach segments, and low correlations between final approach and touchdown scores.

One internal standard that might be employed to evaluate vertical glideslope RMS error for visual system parameters is the difference in RMS error obtained under the two levels of turbulence. The differences in RMS errors between two levels of an equipment factor can be compared with the differences between high and low turbulence levels since these latter levels were chosen to cover an approximate maximum performance range, with no turbulence the easiest condition and high turbulence being set near the maximum level operationally acceptable at sea.

Another standard against which to evaluate the magnitude of the RMS error of the equipment parameters can be based on the values obtained by pilots with different degrees of experience. On average, with no turbulence, the RMS vertical error on final approach for experienced pilots was around 0.3° or less (i.e., within plus or minus one meatball of the FLOLS display most of the time), between 0.3° and 0.55° for novice pilots at the end of training, and greater than 0.55° for novice pilots at the beginning of training. (Novice pilot data was obtained from a later VTRS experiment reported in Westra, 1982.) It is generally assumed that by staying within 0.3° of the glideslope, the final approach will be of high quality. With turbulence as defined in these experiments, these values are all increased by 0.15°. The difference between a highly experienced pilot and a novice was therefore greater than 0.25° of vertical RMS error.

SECTION III

EXPERIMENT I

Experiment I involved the carrier landing task with straight-in approaches. The effects of nine equipment factors were investigated: FLOLS, ship detail, field of view, visual lags, seascape, brightness, line rate, motion and engine lags. In addition, turbulence was also varied. The ship type was a computer-generated image and the G-seat was turned off. The levels of the manipulated factors are given in Table 1.

DESIGN

The basic experimental design is a modified version of the 2^{11-4} fractional factorial plan described in the National Bureau of Standards (1957) report, design 16.11.32. This design includes 128 of the possible 2048 conditions of a full 2^{11} factorial, divided into four blocks of 32 conditions each. The defining generators for the basic design are:

$$I = ABCDJK = ABEFJL = BCEGJKL = ABCDEFGH$$

for which the generalized interactions yield 15 defining contrasts (see Box and Hunter, 1961; Davies, 1967; Simon, 1973; 1977).

The factors--turbulence, seascape, FLOLS, line rate, field of view, ship detail, engine lags, visual lags, motion and brightness--were associated with letters A, B, C, D, E, F, H, J, K, and L respectively. This original design was modified to reduce the chance of pilot fatigue by further dividing each block of 32 conditions in half, basing the division on the high and low levels of Factor G. A different pilot was associated with each of the eight blocks in this final design and was tested on the 16 conditions specific to the block. Block effects in the modified design are totally confounded with the following effects:

$$\text{Blocks} = DEFG = BCFG = BCDE = G = DEF = BCF = BCDEG.$$

Still all main and all two-factor interactions (except H x L)* are independent (orthogonal) of each other and blocks (pilots). Mean differences among pilots, as represented by the mean differences among blocks, can therefore be isolated, but since pilot characteristics were never dimensionalized (assuming differences exist), no interpretation was made of these data in depth.

Each pilot executed eight experimental conditions in the morning and eight more in the afternoon. The experimental conditions were presented to each pilot in an order that kept the estimated effects of the experimental factors robust to potential linear, quadratic, and cubic trend effects (see

---

* The interaction of engine lags with brightness was purposefully chosen to be the H x L interaction and confounded with blocks since it was believed to be one of the two-factor interactions least likely to be important.

Simon, 1977). Counterbalancing across pilots was employed in the case of two factors: line rate and ship detail, in order to further minimize potential trend confounding.

Each experimental condition was run for two consecutive trials. This replication doubled the size of the experiment, i.e., an N of 256 rather than 128. Replication is not ordinarily in keeping with the strategy of economy, but it was done here to increase the power of the experiment. Preliminary studies had suggested that many of the factors being studied would have small or trivial effects on performance. As a number of these factors had large cost implications, it was considered important to be able to state with a high degree of confidence that no meaningful difference existed when this was suspected from the given data. The large sample size was needed to decrease the probability of being wrong when stating that certain factors had no practical effects.

RESULTS

Analyses of the effects of the nine simulator factors plus turbulence on glideslope, lineup and angle of attack RMS error were made for three segments of the descent path (6000'-4500'; 4500'-2000'; 2000'-500'). For the same measures, percent time within specified tolerances was analyzed for the segment 1000' to the ramp. The results are given in Tables 3, 4, and 5. Similar analyses were performed on the wire trapped, lineup error, vertical velocity, pitch, and roll at touchdown, both for the capture values (Table 7) and the quality scores (Table 6). Other results of interest are given in Table 8. The effects of the simulator factors on safe successful landings were also determined.

GENERAL DISCUSSION OF RESULTS. Two-hundred and fifty-six simulated straight-in carrier landings were attempted by experienced pilots. Fifty-nine percent of those attempts resulted in a "successful" landing. The results discussed here will cover performance during the straight-in descent to the carrier ramp and at the point of touchdown.

Descent Phase. When the mean differences in RMS error between levels of the simulator factors are examined (Table 3, 4, and 5), several general results can be observed. First, since the values are expressed in one-hundredth of degrees or units, a value of 10 would be needed to show a one-tenth of a degree or unit difference, an amount that would not ordinarily be considered critical for glideslope and insignificant for lineup and angle of attack. In these tables few values associated with simulator factors exceeded that magnitude of difference.

Second, the large variations in performance during the descent phase were due primarily to turbulence and to the variability among experimental blocks, this latter effect probably reflecting differences among pilots. Only a few large differences were due to variations in simulator factors. The percentage of total performance variability accounted for by the simulator factors further illustrates the small contribution simulator factors played in this task with these subjects. For glideslope, lineup, and angle of attack RMS

TABLE 3.. EXPERIMENT I: GLIDESLOPE PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01°) 6000'-4500'	4500'-2000'	2000'-500'	% time ±.3° 1000'-ramp
FLOLS	Opti-cal	CIG	1	1(-)[2]	6(3.7)*	8(4.8)*	-9(2.5)*
SHIP DETAIL	Day	Night	1	-1(-)	-4(1.4)	-3(-)	5(-)
SEASCAPE	On	Off	1	-3(1.3)	-3(1.2)	2(-)	-5(-)
LINE RATE	1025	525	1	2(-)	-2(-)	0(-)	0(-)
MOTION	On	Off	1	3(1.1)	2(-)	1(-)	0(-)
ENGINE LAGS	30 Hz	7.5 Hz	1	-1(-)	1(-)	-2(-)	6(1.1)
FIELD OF VIEW	Wide	Nar-row	1	-1(-)	-1(-)	1(-)	-1(-)
BRIGHTNESS	2.8 fL	.4 fL	1	0(-)	1(-)	-1(-)	3(-)
VIS. LAGS	100 msec	200 msec	1	-2(-)	-2(-)	-1(-)	2('
TURBULENCE	None	Winds	1	0(-)	-6(3.1)*	-8(5.2·*	12(4.4)*
2 FACTOR INTERACTIONS			44	(19.)	(15.)	(10.)	(10.)
BLOCKS (PILOTS)			7	(22.)*	(18.)*	(11.)	(18.)*
RESIDUAL			194	(54.)	(56.)	(66.)	(62.)
	GRAND MEAN			20	28	30	61
	STD ERR DIFF			1.4	1.7	2.0	3.2

[1]Mean of observations taken under high level minus mean of observations taken under low level of a factor.

[2]Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .005$

TABLE 4. EXPERIMENT I: LINEUP PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01°) 6000'-4500'	4500'-2000'	2000'-500'	% time ±1.5° 1000'-ramp
FLOLS	Optical	CIG	1	1(-) [2]	8(2.0)	6(-)	1(-)
SHIP DETAIL	Day	Night	1	0(-)	-10(3.0)*	-6(-)	-2(-)
SEASCAPE	On	Off	1	1(-)	-6(-)	-5(-)	-1(-)
LINE RATE	1025	525	1	2(-)	5(-)	1(-)	-2(-)
MOTION	On	Off	1	1(-)	5(-)	1(-)	0(-)
ENGINE LAGS	30 Hz	7.5 Hz	1	1(-)	-2(-)	-3(-)	2(-)
FIELD OF VIEW	Wide	Narrow	1	1(-)	0(-)	-4(-)	0(-)
BRIGHTNESS	2.8 fL	.4 fL	1	-1(-)	-2(-)	0(-)	-1(-)
VIS. LAGS	100 msec	200 msec	1	-2(-)	1(-)	-4(-)	0(-)
TURBULENCE	None	Winds	1	-2(-)	-14(5.3)*	-17(4.9)*	3(-)
2 FACTOR INTERACTIONS			44	(17.)	(21.)	(19.)	(12.)
BLOCK (PILOTS)			7	(14.)*	(7.0)*	(13.)*	(8.5)*
RESIDUAL			194	(66.)	(60.)	(61.)	(76.)
	GRAND MEAN			23	41	56	93
	STD ERR DIFF			1.5	3.4	4.3	2.4

[1] Mean of observations taken under high level minus mean of observations taken under low level of a factor.

[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .005$

TABLE 5. EXPERIMENT I: ANGLE OF ATTACK PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01 units) 6000'-4500'	4500'-2000'	2000'-500'	% time ±1 un 1000'-ramp
FLOLS	Optical	CIG	1	-6(-)[2]	-9(1.6)	-6(-)	0(-)
SHIP DETAIL	Day	Night	1	2(-)	-4(-)	-3(-)	6(1.3)
SEASCAPE	On	Off	1	2(-)	0(-)	0(-)	1(-)
LINE RATE	1025	525	1	4(-)	1(-)	1(-)	-1(-)
MOTION	On	Off	1	13(1.6)	3(-)	9(-)	0(-)
ENGINE LAGS	30 Hz	7.5 Hz	1	0(-)	1(-)	-6(-)	2(-)
FIELD OF VIEW	Wide	Narrow	1	-6(-)	-4(-)	2(-)	1(-)
BRIGHTNESS	2.8 fL	.4 fL	1	5(-)	6(-)	-1(-)	4(-)
VIS. LAGS	100 msec	200 msec	1	-7(-)	-7(-)	-2(-)	-2(-)
TURBULENCE	None	Winds	1	-13(1.6)	-30(12.)*	-33(12.)*	14(8.6)*
2 FACTOR INTERACTIONS			44	(13.)	(11.)	(12.)	(14.)
BLOCKS (PILOTS)			7	(27.)*	(20.)*	(20.)*	(16.)*
RESIDUAL			194	(55.)	(53.)	(54.)	(57.)
	GRAND MEAN			94	101	116	54
	STD ERR DIFF			5.5	4.6	5.1	2.8

[1] Mean of observations taken under high level minus mean of observations taken under low level of a factor.

[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .005$

TABLE 6. EXPERIMENT I: TOUCHDOWN QUALITY EFFECTS

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] Wire Trapped	Lineup	Vert. Vel	Pitch	Roll
FLOLS	Opti-cal	CIG	1	-1.3(-)[2]	3.4(-)	-2.6(-)	1.7(-)	4.8(-)
SHIP DETAIL	Day	Night	1	7.3(1.0)	5.5(-)	3.2(-)	-2.6(-)	6.4(1.1)
SEASCAPE	On	Off	1	3.7(-)	0.3(-)	1.2(-)	-0.7(-)	1.3(-)
LINE RATE	1025	525	1	-7.0(-)	-1.7(-)	-4.7(-)	-8.5(1.3)	-1.3(-)
MOTION	On	Off	1	-5.5(-)	-3.7(-)	5.1(-)	-10.9(2.0)	-3.0(-)
ENGINE LAGS	30 Hz	7.5 Hz	1	-1.5(-)	13.3(3.0)*	7.9(1.1)	9.1(1.5)	7.0(1.4)
FIELD OF VIEW	Wide	Nar-row	1	-0.2(-)	2.6(-)	7.3(-)	5.8(-)	7.5(1.5)
BRIGHTNESS	2.8 fL	.4 fL	1	-2.4(-)	-6.7(-)	6.5(-)	-1.6(-)	5.2(-)
VIS. LAGS	100 msec	200 msec	1	3.0(-)	-2.3(-)	-0.4(-)	-1.4(-)	3.5(-)
TURBULENCE	None	Winds	1	4.3(-)	6.8(-)	0.9(-)	7.0(-)	6.8(1.3)
2 FACTOR INTERACTIONS			44	(11.)	(18.)	(15.)	(15.)	(15.)
BLOCKS (PILOTS)			7	(17.)*	(15.)*	(3.6)	(4.2)	(13.)*
RESIDUAL			194	(69.)	(62.)	(78.)	(74.)	(65.)
			GRAND MEAN	54.3	67.9	65.5	64.6	78.5
			STD ERR DIFF	4.3	4.3	4.7	4.5	3.5

[1]Mean of observations taken under high level minus mean of observations taken under low level of a factor expressed in quality points.

[2]Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .005$

TABLE 7. EXPERIMENT I: MEANS AND DISPERSIONS AT TOUCHDOWN

FACTOR		Wire Trapped[1] high	Wire Trapped[1] low	Lineup Error (ft) high	Lineup Error (ft) low	Vert. Vel. (ft/sec) high	Vert. Vel. (ft/sec) low	Pitch (degrees) high	Pitch (degrees) low	Roll (degrees) high	Roll (degrees) low	Sample Size[2] high	Sample Size[2] lo[illegible]
FLOLS	Mean	176	160	2.7	-.3	8.2	8.8	8.1	8.0	-.4	-.9		
	Disp	89	67	9.4	10.7	2.2	2.6	2.0	2.1	2.8	3.1	119	11[illegible]
SHIP DETAIL	Mean	183	152*	0.0	2.6	8.7	8.3	7.8	8.2	-1.0	-.3		
	Disp	83	73	10.5	9.7	2.5	2.3	2.1	1.9	2.9	3.0	123	11[illegible]
SEASCAPE	Mean	176	160	0.7	1.8	8.9	8.2	7.8	8.2	-.8	-.5		
	Disp	81	78	10.3	10.0	2.4	2.4	2.1	2.0	3.1	2.8	118	11[illegible]
LINE RATE	Mean	180	156	3.1	-.6*	8.6	8.4	8.0	8.0	-.5	-.8		
	Disp	90	65	10.2	9.8	2.4	2.4	2.2	1.9	2.9	3.0	117	11[illegible]
MOTION	Mean	163	172	.1	2.4	8.9	8.1	7.6	8.5*	-.9	-.4		
	Disp	77	82	9.4	10.8	2.0	2.7	2.0	1.9	2.8	3.1	114	12[illegible]
ENGINE LAGS	Mean	169	167	.6	2.0	8.2	8.8	8.1	7.9	-.9	-.3		
	Disp	78	81	9.0	11.2	2.3	2.5	2.0	2.1	2.8	3.1	123	11[illegible]
FIELD OF VIEW	Mean	173	163	1.0	1.5	8.6	8.4	8.1	7.9	-1.0	-.3		
	Disp	86	72	10.6	9.6	2.3	2.5	2.1	2.0	2.7	3.2	120	11[illegible]
BRIGHTNESS	Mean	166	170	1.0	1.5	8.5	8.6	8.0	8.1	-.7	-.6		
	Disp	79	80	9.9	10.4	2.4	2.4	2.1	2.0	2.8	3.1	121	11[illegible]
VISUAL LAGS	Mean	166	170	.8	1.7	8.6	8.4	8.1	8.0	-.5	-.8		
	Disp	83	76	10.0	10.3	2.4	2.4	1.9	2.1	2.6	3.3	117	11[illegible]
TURBULENCE	Mean	159	178	1.3	1.2	8.1	9.0*	8.5	7.6*	-.5	-.8		
	Disp	70	87	9.6	10.7	2.3	2.4	1.9	2.0	2.8	3.2	121	11[illegible]

[1] Values shown are for distance from the ramp in feet. The #3 wire is at 186 feet and the distance between the wires is approximately 33 feet.
[2] Refers to the number of trials that terminated in a touchdown on the landing deck.
* $p \leq .005$ for mean difference between levels

TABLE 8. EXPERIMENT I: SUMMARY OF SELECTED PERFORMANCE MEASURES

SOURCE OF VARIANCE	LEVELS HIGH	LEVELS LOW	df	MEAN DIFFERENCE[1] LINEUP BIAS 6000'-4500' (degrees)	AOA BIAS 4500'-2000' (units)	ROLL VAR 4500'-2000' (degrees)	ROLL VAR 1000'-Ramp (degrees)	PITCH VAR 1000'-Ramp (degrees)	ROLL BIAS 1000'-Ramp (degrees)
FLOLS	Optical	CIG	1	-.01(-)[2]	-.06(-)	.00(-)	-.14(-)	.07(-)	-.11(-)
SHIP DETAIL	DAY	NIGHT	1	.18(14.)*	.07(-)	-.17(-)*	-.09(-)	-.04(-)	.09(-)
SEASCAPE	ON	OFF	1	.05(1.1)*	.09(-)	.00(-)	-.10(-)	.07(-)	.15(1.3)
LINE RATE	1025	525	1	.00(-)	.14(-)	-.03(-)	.04(-)	.01(-)	.27(4.2)*
MOTION	ON	OFF	1	-.06(1.4)	-.31(4.3)*	.09(-)	.04(-)	.03(-)	.00(-)
ENGINE LAGS	30 Hz	7.5 Hz	1	.02(-)	-.01(-)	.03(-)	-.15(-)	-.16(2.5)*	.09(-)
FIELD OF VIEW	Wide	Narrow	1	.02(-)	.04(-)	-.33(4.5)*	-.44(43)*	.08(-)	.22(2.9)*
BRIGHTNESS	2.8 fL	.4 fL	1	.01(-)	.09(-)	-.13(-)	-.28(1.7)*	-.01(-)	-.03(-)
VISUAL LAGS	100 msec	200 msec	1	.02(-)	.08(-)	-.27(2.9)*	-.45(4.5)*	-.02	-.00 (-)
TURBULENCE	None	Winds	1	.02(-)	.20(1.8)	-.80(25.)*	-.96(21.)*	-.53(27.)*	-.15(1.3)
2 FACTOR INTERACTIONS			44	(12.)	(12.)	(6.6)	(7.6)	(12.)	(13.)
BLOCKS (PILOTS)			7	(26.)*	(30.)*	(33.)*	(25)*	(17)*	(24.)*
RESIDUAL			194	(44.)	(50.)	(24.)	(34.)	(40.)	(52.)
GRAND MEAN				.06	.01	1.61	2.21	1.07	-.14
STD ERR DIFF				.023	.076	.058	.089	.046	.068

[1] Mean of observations taken under high level minus mean of observations taken under low level of a factor.

[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .005$

error, the percent of total variance accounted for by all main simulator effects (but not turbulence) was considerably less than 10% across the board. If the percentages accounted for by the two-factor interactions are added, statistically significant or not, the amount increases to around 20%. More than half of the observed variability in performance (as reflected by the residual) is unaccounted for.

Third, across the three tables, the two-factor interactions do not appear to have a significant effect on performance during the descent phase.

Touchdown Phase. Similar general results can be observed in Tables 6 and 7 containing touchdown data. There were few outstanding effects due to the simulator factors. Most of the variations in performance came from block (or pilot) differences. Even the effects of turbulence are less evident at touchdown.

DISCUSSION OF INDIVIDUAL FACTORS. The effects of each simulator factor on various aspects of the carrier landing are discussed below.

FLOLS. The CIG FLOLS results in better performance, on average, than the OPTICAL FLOLS. There was a smaller vertical RMS error along the glideslope when the CIG FLOLS was used, at least from 4500' to 500' (see Table 3). The difference was small but statistically significant at the .005 probability level: 0.25° average RMS error for the CIG FLOLS and 0.32° average RMS error for the Optical FLOLS, with a standard error of the mean difference of ±.018°. The effects of the FLOLS on lineup and angle of attack RMS error and other performance measures were small and inconsistent (Tables 4, 5 and 8) as to be expected since it only provided vertical position information.

At touchdown, with both systems, landings occurred on average between the two and three wires; however, the longitudinal dispersion of touchdown points was greater for the Optical FLOLS (Table 7). No difference in wire- trapped touchdown quality between the two systems could be reliably determined (Table 6). There was no signficant difference between the two FLOLS systems in terms of the number of "successful" touchdowns made with each. With CIG FLOLS, 61% of the landings were successful; with the Optical FLOLS, 58% were successful.

Ship Detail. The high-detail day ship configuration produced better lineup performance than the low-detail night ship configuration. The RMS error (Table 4) and bias error for line-up are both larger for the low-detail ship in the 4500'-2000' segment. (A positive lineup bias error indicates error to the left of centerline.) However, these numbers do not tell the whole story which is better understood by referring to Figure 3. This figure shows that on average when the high-detail ship was used, pilots tracked the centerline with virtually no bias error from 3000 ft. to the ramp while with low ship detail there was a consistent right of centerline bias for this segment. It should be noted that because of ship movement, there will be a tendency to drift left if there is no heading compensation or right wing down correction for left bias. It appears this normal tendency took place in the early part of the approach under high-ship detail and that by 3000 ft. the appropriate

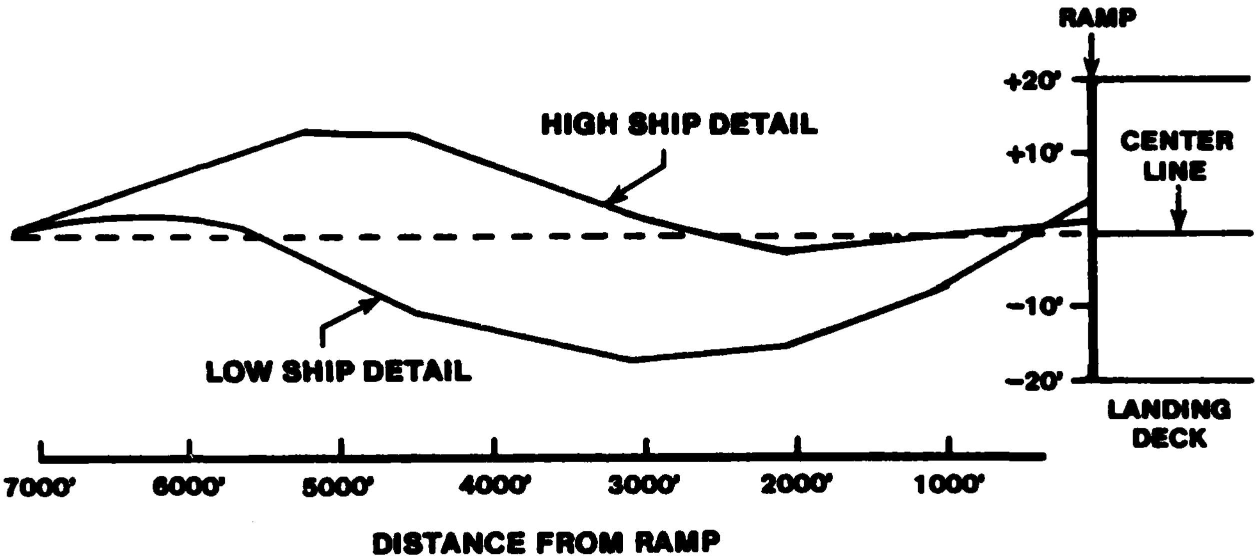


Figure 3. Illustration of Ship Detail Lineup Bias Effect.

corrective action had been taken. However, under low ship detail, presumably because lineup position could not be seen as clearly, pilots overcorrected the tendency to drift left by going too far to the right. This effect was quite consistent across pilots. This maneuver was also reflected in a difference in roll variability for the two ship detail configurations (Table 8). In spite of the statistically significant differences, both flight paths fell well within the 1.5° tolerance band defining acceptable lineup performance.

Ship detail also affected how early the touchdown occurred. On the average, the landing was just in front of the two wire when the low-detail ship was displayed, and just before the three wire when the high-detail ship was used. While statistically significant, both conditions fall within acceptable limits (Table 7, wire-trapped).

Fifty-eight percent of the landings were successful when the low-detail ship was displayed. Sixty-one percent were successful when the high-detail ship was used. This difference is not statistically significant.

Other Visual Factors. Field of view, Brightness, and Visual Lags all show strong effects on the roll variability of the aircraft between 4500' and the ramp (Table 8). In all cases, the high levels (i.e., wide screen, high brightness, and shorter lags) produced significantly smaller roll error variability. The high levels of visual lag and field of view also produced smaller roll dispersions at touchdown (Table 7). Field of view and line rate also showed a significant difference between the two levels for roll bias in the last 1000' to the ramp, with low levels (525 line rate and narrow screen) resulting in left wing down biases. None of these measures, however, appeared to affect RMS error during the descent phase (Tables 3, 4, 5, and 6). Seascape had no effects during the descent. Seascape and line rate were the only two simulator factors in Experiment I that appeared to affect the percentage of successful landings that were made (albeit marginally). With seascape absent, 54% successful landings were made; with it present, 65% were made. The difference is statistically significant at the $p = .08$ level and cannot be taken too seriously without confirmation from other sources.

With the 525 line rate, there were 66% successful landings; with 1025 line rate, there were 52%. This difference in a restricted analysis of variance of Experiment I data was statistically significant at around the $p = .025$ level. Still line rate accounted for only two percent of the total variance in the experiment fcr this measure of successful landings. The surprising reversal of results from what was expected is discussed below.

Table 8 indicates a roll bias effect for line rate in the 1000' to the ramp segment. A more detailed examination of this effect revealed that at 1000' from the ramp the average lineup was approximately 5' to the right of centerline under both 525 and 1025 line rates. However, under the 1025 line rate, the average roll at 1000' was -.76° and the average lineup at the ramp was 6.0' to the left of centerline with an average roll of 1.3°. (Positive roll indicates right wing down.) This compares to an average roll of -.66° at 1000' for the 525 line rate with average lineup at the ramp 0.6' to the left of centerline and average roll of 0.1°. Apparently, with the 1025 line rate

an overcorrection for lineup was made at 1000' necessitating a large countercorrection at the ramp. This lineup problem under 1025 probably is responsible for the poorer landing success rate seen for the 1025 line rate. Why this result occurred in the face of expectations of better performance with the 1025 line rate is not apparent at this time. The effect was somewhat inconsistent across pilots and, although no statistically significant interactions were noted, occurred mostly under low-ship detail; the condition which resulted in a right of centerline error bias at 1000'.

Non Visual Factors. The effects of Engine Lags and Motion were marginal in this experiment. Several isolated effects were statistically significant, but cannot be verified as being meaningful to any practical extent. Engine lags showed an effect at touchdown in the quality of lineup (Table 6) and it also showed a large effect on pitch variability within 1000' of the ramp (Table 8). In both cases, the slower update resulted in poorer performance. There were 61% successful landings with the 30 Hz engines compared to 58% for the 7.5 Hz engines. With motion on, there was a tendency to fly a lower angle of attack (faster airspeed) between 4500' - 2000' than with the no-motion condition (Table 8). This agrees with the observation that the mean pitch of the aircraft at touchdown was lower with the motion condition. Both conditions however are well within the acceptable tolerance limits for these parameters. Under the motion condition, 56% of the landings were successful. Under the no-motion condition, 63% were successful. This difference is not statistically significant.

Other Sources of Variance. The collective estimate of two-factor interaction effects failed to show a statistically significant effect on the percent of successful landings, the p value being greater than .50.

Pilots showed a large statistically significant difference ($p = .005$) in the percent of successful landings each made. It is noteworthy that, on the average, supposedly homogeneous pilots varied considerably more than the purposely varied simulator factors. Out of 32 trials, the number of successful landings ranged from 28 (88%) for one pilot to 12 (38%) for two others. However, since pilot effects are actually block effects, the experimental design enables the effects of other factors to be isolated without being confounded by average pilot differences.

SECTION IV

EXPERIMENT II

This experiment involved the carrier-landing task with circling approaches. THe flight path for the carrier landing mission included a final turn which commenced approximately abeam of the LSO platform on the downwind leg of the circling approach (see Figure 1). The factors varied in this experiment were a subset of the factors included in Experiment I and were selected on the basis of their potential effect on this final turn. Five equipment factors were investigated: ship detail, seascape, motion, brightness, and visual lags. In addition, turbulence was also varied. See Table 1 for a description of the levels used. Factors held constant were field of view (wide), FLOLS (CIG), engine lags (30 Hz update), line rate (1025), ship type (CIG), and G-seat (off).

DESIGN

The basic experimental design was a full 2^6 factorial requiring 64 conditions, divided into four blocks of 16 conditions each. Two pilots were assigned to each block and were tested on two consecutive trials for each of the 16 conditions in the block, bringing the total number of trials in the experiment to 256. All main, two-factor interaction, and block effects are independent of each other and confounded only with three-way and higher-order sources of variance which consist of interactions involving blocks. The design was counterbalanced across pilots to balance main effects against trends.

RESULTS

The analyses performed in Experiment I were performed in Experiment II on the same performance measures. With the introduction of the turn maneuver for the circular approach, new measures were added: Distance from the ramp at the 90° point of the turn, and lineup error at rollout. Other measures were also examined, but none yielded additional information about factor effects. Because of the inclusion of the turn, the "FLOLS space entry to Rollout" segment was analyzed in lieu of the 6000' - 4500' and 4500' - 2000' segments used in the straight-in study. FLOLS space entry was defined as the point at which the pilot can first obtain glideslope information from the FLOLS (see Figure 1) and rollout was defined as the point at which the aircraft heading has crossed the ship's heading and aircraft roll was within ±4° of level.

In the analyses of variance, the sources of variance were partitioned into the effects of the individual factors, two-factor interactions, blocks (pilots), and a residual. Effects (i.e., mean differences between the two-factor levels), statistical significance of the differences, and proportion of total variance accounted for by each source were calculated. A $p < .01$ level was considered reasonable in this experiment to mark the noteworthy sources of variance.

The results from the analyses in Experiment II are shown in Tables 9 through 14.

TABLE 9. EXPERIMENT II: GLIDESLOPE PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01°) TURN[2]	MEAN DIFFERENCE[1] RMS Errors (in .01°) 2000'-500'	MEAN DIFFERENCE[1] %time ± .3° 1000'-ramp
VISUAL LAGS	100 msec	200 msec	1	-2(-)[3]	2(-)	2(-)
BRIGHTNESS	2.8 fL	0.4 fL	1	-4(1.2)	-3(-)	-1(-)
MOTION	On	Off	1	-1(-)	-1(-)	-2(-)
SEASCAPE	On	Off	1	-0(-)	-3(-)	1(-)
SHIP DETAIL	Day	Night	1	0(-)	0(-)	0(-)
TURBULENCE	None	Winds	1	-10(6.3)*	-12(12.)*	21(15.)*
2 FACTOR INTERACTIONS			15	(5.4)	(5.8)	(3.3)
BLOCKS (PILOTS)			7	(13.)*	(14.)*	(12.)*
RESIDUAL			227	(74.)	(67.)	(68.)
	GRAND MEAN			33	28	59
	STD ERR DIFF			2.3	1.9	3.0

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.
[2] The final turn from entry into FLOLS space to rollout for final approach.
[3] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).
* $p \leq .01$

TABLE 10. EXPERIMENT II: LINEUP PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01°) 2000'-500'	MEAN DIFFERENCE[1] % time ± 1.5° 1000'-ramp
VISUAL LAGS	100 msec	200 msec	1	-2(-)[2]	-1(-)
BRIGHTNESS	2.8 fL	0.4 fL	1	4(-)	-4(1.3)
MOTION	On	Off	1	5(-)	0(-)
SEASCAPE	On	Off	1	-2(-)	3(-)
SHIP DETAIL	Day	Night	1	2(-)	1(-)
TURBULENCE	None	Winds	1	-2(-)	1(-)
2 FACTOR INTERACTIONS			15	(8.4)	(9.7)
BLOCKS (PILOTS)			7	(18.)*	(9.8)*
RESIDUAL			227	(72.)	(77.)
GRAND MEAN				56	88
STD ERR DIFF				3.6	2.1

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.

[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .01$

TABLE 11. EXPERIMENT II: ANGLE OF ATTACK PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01 units) TURN[2]	MEAN DIFFERENCE[1] RMS Errors (in .01 units) 2000'-500'	MEAN DIFFERENCE[1] % time ± 1 un 1000-ramp
VISUAL LAGS	100 msec	200 msec	1	-4(-)[3]	-7(-)	0(-)
BRIGHTNESS	2.8 fL	0.4 fL	1	1(-)	1(-)	-3(-)
MOTION	On	Off	1	7(-)	4(-)	-4(-)
SEASCAPE	On	Off	1	0(-)	2(-)	-3(-)
SHIP DETAIL	Day	Night	1	0(-)	2(-)	-1(-)
TURBULENCE	None	Winds	1	-22(4.1)*	-47(19.)*	10(12.)*
2 FACTOR INTERACTIONS			15	(3.1)	(4.5)	(6.6)
BLOCKS (PILOTS)			7	(46.)*	(13.)*	(2.7)
RESIDUAL			227	(46.)	(62.)	(77.)
	GRAND MEAN			123	127	48
	STD ERR DIFF			4.9	5.6	1.7

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.
[2] The final turn from entry into FLOLS space to rollout for final approach.
[3] Values in parenthesis are percent variance accounted for in the experiment. Percent less than 1.0 are shown by a dash (-).
* $p \leq .01$

TABLE 12. EXPERIMENT II: TOUCHDOWN QUALITY EFFECTS

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] Wire Trapped	Lineup	Vert. Vel	Pitch	Roll
VIS. LAGS	100 msec	200 msec	1	4.3(-)[2]	3.9(-)	4.9(-)	-4.7(-)	5.4(1.0)
BRIGHTNESS	2.8 fL	0.4 fL	1	-.5(-)	-2.5(-)	7.2(1.0)	-2.9(-)	-6.6(1.5)
MOTION	On	Off	1	3.2(-)	11.3(2.2)*	9.6(1.7)	6.7(-)	5.8(1.2)
SEASCAPE	On	Off	1	3.4(-)	1.4(-)	-9.5(1.4)	-7.4(-)	0.5(-)
SHIP DETAIL	Day	Night	1	12.0(2.3)*	7.3(-)	-1.0(-)	5.6(-)	9.6(3.2)*
TURBULENCE	None	Winds	1	25.2(11.)*	8.4(1.5)	20.4(6.9)*	16.5(4.7)*	14.1(7.2)*
2 FACTOR INTERACTIONS			15	(8.2)	(9.2)	(10.)	(6.0)	(7.2)
BLOCKS (PILOTS)			7	(5.6)*	(13.)*	(0.6)	(4.9)	(6.5)*
RESIDUAL			227	(74.)	(74.)	(78.)	(83.)	(71.)
	GRAND MEAN			66.3	69.1	65.4	62.6	84.4
	STD ERR DIFF			4.3	4.3	4.5	4.5	2.9

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor expressed in quality points.

[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .01$

TABLE 13. EXPERIMENT II: MEANS AND DISPERSIONS AT TOUCHDOWN

FACTOR		Wire Trapped[1] High	Wire Trapped[1] Low	Lineup Error (ft) High	Lineup Error (ft) Low	Vert. Vel. (ft/sec) High	Vert. Vel. (ft/sec) Low	Pitch (degrees) High	Pitch (degrees) Low	Roll (degrees) High	Roll (degrees) Low	Sample Size[2] High	Sample Size[2] Low
VISUAL LAGS	Mean	172	166	1.6	2.4	8.9	9.1	7.5	7.6	-0.2	-0.1		
	Disp	63	75	9.1	10.3	2.6	3.5	2.1	2.6	2.0	2.8	119	121
BRIGHTNESS	Mean	174	165	2.1	1.9	9.5	8.6*	7.3	7.8	0.0	-0.2		
	Disp	70	68	10.3	9.0	2.9	3.3	2.4	2.4	2.2	2.7	122	118
MOTION	Mean	177	161	2.2	1.8	8.8	9.3	7.5	7.5	-0.1	-0.1		
	Disp	74	64	8.8	10.5	2.9	3.2	2.4	2.4	2.3	2.6	123	117
SEASCAPE	Mean	175	163	1.3	2.7	9.5	8.6	7.2	7.9	-.2	0.0		
	Disp	68	71	9.7	9.6	3.4	2.7	2.5	2.2	2.4	2.5	120	120
SHIP DETAIL	Mean	169	170	0.9	3.1	9.3	8.7	7.3	7.8	-0.5	0.3		
	Disp	61	78	9.3	10.0	3.0	3.1	2.6	2.1	2.1	2.7	124	116
TURBULENCE	Mean	177	161	1.7	2.3	8.3	9.8*	7.6	7.5	-0.1	-0.1		
	Disp	55	81	9.1	10.3	2.3	3.6	2.0	2.7	2.1	2.8	126	114

[1]Values shown are for distance from the ramp in feet. The #3 wire is at 186 feet and the distance between the wires is approximately 33 feet.

[2]Refers to the number of trials that terminated in a touchdown on the landing deck.

* $p \leq .01$ for mean difference between levels

TABLE 14. EXPERIMENT II: SUMMARY OF SELECTED PERFORMANCE MEASURES

SOURCE OF VARIANCE	LEVELS HIGH	LEVELS LOW	df	MEAN DIFFERENCE[1]: AOA ERROR BIAS 1000-RAMP (units)	DISTANCE AT 90° (feet)	ROLL AT 90° (degrees)	LINEUP ERROR AT ROLLOUT (feet)	ROLL VAR. 1000'-RAMP (degrees)
VISUAL LAGS	100 msec	200 msec	1	-.07(-)[2]	-41(-)	-.33(-)	9.2(-)	-.43(4.0)*
BRIGHT-NESS	2.8 fL	0.4 fL	1	.09(-)	-109(-)	-.18(-)	-4.0(-)	-.13(-)
MOTION	ON	OFF	1	-.20(1.2)	88(-)	-.59(1.5)	23.7(1.6)	.04(-)
SEASCAPE	ON	OFF	1	-.37(3.5)*	-256(4.0)*	.04(-)	8.6(-)	.02(-)
SHIP DETAIL	DAY	NIGHT	1	-.01(-)	72(-)	-1.48(8.7)*	46.0(5.9)*	-.02(-)
TURBU-LENCE	None	Winds	1	-.27(2.3)*	44(-)	-.78(2.7)*	5.2(-)	-.97(21.)*
_ FACTOR INTERACTIONS			15	(2.6)	(2.2)	(3.3)	(5.4)	(5.8)*
BLOCK (PILOTS)			7	(23.)*	(30.)*	(18.)*	(27.)*	(27.)*
RESIDUAL			227	(67.)	(63.)	(65.)	(59.)	(42.)
		GRAND MEAN		-.77	4997	-11.6	67.6	2.5
		STD ERR DIFF		.11	67.4	.27	9.6	.09

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.
[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).
* $p \leq .01$

GENERAL DISCUSSION OF RESULTS. Two-hundred and fifty-six simulated circular carrier landings were made by eight experienced pilots, all having had previous simulator experience during Experiment I but with straight-in landings. Sixty-three percent of these were successful, four percent more than with straight-in landings. Under the assumption that the circling task was more difficult, this may reflect learning, but it is also true that not all combinations of difficult conditions were involved in this second study, particularly the half containing the poorer Optical FLOLS configuration.

Tables 9, 11, 12 and 14 reveal that turbulence and blocks (reflecting pilot differences) again account for most of the definable variability in the experiment. Turbulence particularly showed a much larger effect on the quality of touchdown (Table 12) with the circling approach than it had with the straight-in approach to landing. However, turbulence showed no effect on lineup error at any point during the descent (Table 10), unlike the results in Experiment I (Table 4). This might be explained by the pilots' increased ability to handle turbulence, in general, achieving faster learning in the lateral dimensions where turbulence involved lower frequency than in the vertical dimension where the frequencies were higher. In general, the combined two-factor interactions showed no statistically significant effect ($p \leq .01$).

DISCUSSION OF INDIVIDUAL FACTORS. Those individual factors showing an effect on performance during this second (circular approach) experiment are discussed below.

Ship Detail. Because of the circular approach in this experiment, no measure of lineup error was taken beyond 2000 ft. from the ship's ramp (Table 10). Thus, there is no way to compare the effect of ship detail in the 4500' - 2000' segment observed on lineup error in Experiment I (Table 4). But a similar effect does show up during the turn (Table 14) with what appears to be a larger lineup error at rollout when the high-detail ship is used. As in experiment I, however, the in close and touchdown lineup performance (Table 13) is better with the high-detail ship on average. Following rollout, the flight paths are similar to those shown in Figure 3, although there is not as much bias error to the right under low ship detail.

High ship detail resulted in a significantly higher touchdown quality wire trapped score than did low ship detail (Table 12). This may have been the result of less dispersion at landing with that display (Table 13). The low-detail ship design extended the range of landings in such a way that, although the mean touchdown position was essentially the same (Table 13), the greater dispersion resulted in more landings of unacceptable quality. Seventy percent of the attempted landings were within acceptable limits when the high-detail ship configuration was used and only 56 percent of the landings were acceptable with the low-detail ship configurations. The difference was marginally significant ($p = .02$). This difference in levels of ship detail compares with a 28 percent difference for calm (77%) and wind gust (49%) turbulence conditions in Experiment II.

Seascape. The larger angle-of-attack (AOA) bias error (Table 14) indicated that the final approach (1000' to ramp) was flown faster, on the average, with the seascape on than with it off. This may account for the slightly harder landings at touchdown indicated in Table 13 and the vertical velocity touchdown quality effect (albeit marginal) in Table 12. During the turn, Seascape registered a statistically signficant effect (Table 14). Specifically, the mean distance at the 90° point of the turn was 256 feet further from the carrier ramp when the seascape was off (5127 ft.) than when the seascape was on (4871 ft.), with essentially no difference in dispersion. With seascape off, 68% of the landings were successful. With it on, 58% of them were successful. This difference was marginally significant ($p = .07$).

Although these results suggest that performance may have been hurt when seascape was present, the practical implications of this difference are not known. However, it is interesting to note that pilots are instructed not to use the seascape as a cue for distance or altitude at sea.

Other Visual Factors. Brightness showed a marginal effect on roll quality at touchdown (Table 12). Visual lags also showed an effect on roll variability in the segment between 1000' and the ramp (Table 14). These results are consistent with the general effects on roll shown by visual factors in the first experiment.

While there is no difference in the percent of successful landings for the two levels of brightness, visual lags showed a marginal effect ($p = .04$), with 69% of the landings successful when the lag was at 100 msec and 57% when the longer lag (200 msec) was introduced.

MOTION. Motion appeared to have an effect on touchdown lineup quality (Table 12) but the difference in lineup dispersion at touchdown was not large (Table 13). Sixty-five percent of the successful landings occurred with motion on and sixty-one percent with motion off. The difference is not statistically significant ($p = .49$).

SECTION V

EXPERIMENT III

This experiment involved the carrier landing task with straight-in approaches. Two equipment factors, ship type and G-seat, plus turbulence were studied at the levels shown in Table 1. These two equipment variables had been omitted from the earlier experiments because of technical problems. The factors held constant were FLOLS (CIG), ship detail (solid surface), field of view (wide), visual lags (100 msec), seascape (on), brightness (high), line rate (1025), motion (off), and engine lags (30 Hz update).

DESIGN

The basic experimental design was a 2^3 factorial. Each of the eight pilots (used in Experiments I and II) were tested on two consecutive trials on all eight conditions. This resulted in a total of 128 observations. The design was counterbalanced across subjects to minimize trend effects. All experimental factors and interaction effects were estimable.

RESULTS

The analyses performed in Experiment I were also performed on the same performance measures in Experiment III. In the analysis of variance the two main simulator factors, turbulence, and their two-factor interactions were isolated along with block effects. A $p < .05$ was considered a reasonable level for statistical significance in this study. The results from these analyses are shown in Tables 15 through 19.

GENERAL DISCUSSION OF RESULTS. Turbulence and pilots (blocks) were again the major sources of variance for most of the measures in this third experiment. Residual variance - the unexplained variability - was by far the highest as before. The two simulator factors, ship type and G-seat, plus their interaction generally accounted for less than 5% of all variability on any of the different measures used in this experiment.

Of the 128 straight-in trials, 84% of the landings were successful. This was approximately 20% more than in the two previous experiments. While it is probably true that the pilots were learning to fly the simulator better, it should be noted that those factors which appeared to have the largest effects on performance in the previous experiments were included only at their preferred level in this study. This meant that the bulk of the more difficult conditions in Experiments I and II were never experienced in this experiment.

DISCUSSION OF INDIVIDUAL FACTORS. With the exception of turbulence, none of the factors in the first two experiments were included in this third one. The effects of G-seat and ship type, their interaction with each other and with turbulence are discussed below.

TABLE 15. EXPERIMENT III: GLIDESLOPE PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01°) 6000-4500'	4500-2000'	2000-500'	% time ±.3° 1000'-ramp
SHIP TYPE	Model	CIG	1	2(1.0)[2]	2(1.4)	-1(-)	2(-)
G-SEAT	On	Off	1	-2(-)	-1(-)	2(-)	-9(3.4)*
TURBULENCE	None	Winds	1	0(-)	-5(5.0)*	-6(7.6)*	14(9.0)*
SHIP TYPE BY G-SEAT			1	(1.2)	(-)	(1.1)	(1.1)
SHIP TYPE BY TURBULENCE			1	(1.0)	(1.4)	(1.4)	(-)
G-SEAT BY TURBULENCE			1	(-)	(-)	(-)	(-)
BLOCKS (PILOTS)			7	(20.)*	(22.)*	(17.)*	(14.)*
RESIDUAL			114	(76.)	(69.)	(72.)	(72.)
	GRAND MEAN			15	21	21	75
	STD ERR DIFF			1.6	1.7	1.7	3.7

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.
[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).
* $p \leq .05$

TABLE 16. EXPERIMENT III: LINEUP PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] RMS Errors (in .01°) 6000-4500'	4500-2000'	2000-500'	% time ±1.5° 1000'-ramp
SHIP TYPE	Model	CIG	1	-4(2.2)[2]	6(2.7)*	8(2.5)*	1(-)
G-SEAT	On	Off	1	2(-)	8(4.2)*	3(-)	0(-)
TURBULENCE	None	Winds	1	3(1.4)	5(2.2)	7(1.9)	1(-)
SHIP TYPE BY G-SEAT			1	(-)	(3.1)*	(-)	(-)
SHIP TYPE BY TURBULENCE			1	(2.4)*	(1.3)	(1.3)	(-)
G-SEAT BY TURBULENCE			1	(-)	(-)	(-)	(-)
BLOCK (PILOTS)			7	(30.)*	(23.)*	(22.)*	(9.2)
RESIDUAL			114	(64.)	(63.)	(72.)	(89.)
		GRAND MEAN		22	30	43	98
		STD ERR DIFF		2.0	2.9	4.0	-1.6

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.
[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).
* $p \leq .05$

TABLE 17. EXPERIMENT III: ANGLE OF ATTACK PERFORMANCE

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE [1] RMS Errors (in .01 units) 6000-4500'	4500-2000'	2000-500'	% time ±1 unit 1000'-ramp
SHIP TYPE	Model	CIG	1	10(1.2)[2]	2(-)	4(-)	-3(-)
G-SEAT	On	Off	1	4(-)	3(-)	4(-)	-2(-)
TURBULENCE	None	Winds	1	0(-)	-23(9.0)*	-29(12.)*	13(7.9)*
SHIP TYPE BY G-SEAT			1	(-)	(-)	(-)	(-)
SHIP TYPE BY TURBULENCE			1	(-)	(-)	(1.5)	(2.2)*
G-SEAT BY TURBULENCE			1	(-)	(1.9)	(-)	(-)
BLOCKS (PILOTS)			7	(24.)*	(32.)*	(28.)*	(31.)*
RESIDUAL			114	(73.)	(57.)	(57.)	(58.)
		GRAND MEAN		90	94	100	61
		STD ERR DIFF		7.3	5.4	5.9	3.3

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.
[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).
* $p \leq .05$

TABLE 18. EXPERIMENT III: TOUCHDOWN QUALITY EFFECTS

SOURCE OF VARIANCE	LEVELS High	LEVELS Low	df	MEAN DIFFERENCE[1] Wire Trapped	Lineup	Vert. Vel	Pitch	Roll
SHIP TYPE	Model	CIG	1	-2.2(-)[2]	-5.5(-)	5.1(-)	-2.2(-)	-4.7(1.6)
G-SEAT	On	Off	1	2.3(-)	-10.0(3.3)*	5.7(-)	-2.7(-)	-1.2(-)
TURBULENCE	None	Winds	1	-14.8(7.4)*	4.8(-)	3.4(-)	-0.1(-)	6.7(3.2)*
SHIP TYPE BY G-SEAT			1	(-)	(1.6)	(2.3)	(2.9)	(-)
SHIP TYPE BY TURBULENCE			1	(-)	(-)	(3.5)*	(-)	(-)
G-SEAT BY TURBULENCE			1	(5.4)*	(-)	(1.3)	(-)	(-)
BLOCKS (PILOTS)			7	(15.)*	(27.)*	(6.9)	(12.)*	(29.)*
RESIDUAL			114	(71.)	(67.)	(85.)	(84.)	(65.)
	GRAND MEAN			79.6	76.8	75.7	71.3	85.7
	STD ERR DIFF			4.3	4.2	5.1	5.1	2.8

[1] Mean of observations taken under high level minus mean of observations taken under low level of factor.

[2] Values in parenthesis are percent variance accounted for in the experiment. Percents less than 1.0 are shown by a dash (-).

* $p \leq .05$

TABLE 19. EXPERIMENT III: TOUCHDOWN MEANS AND DISPERSIONS

FACTOR		Wire Trapped[1]		Lineup Error (ft)		Vert. Vel. (ft/sec)		Pitch (degrees)		Roll (degrees)		Sample Size[2]	
		High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
SHIP TYPE	Mean	153	168*	-2.8	-2.5	9.4	8.8	7.4	7.8	-1.2	-2.0*		
	Disp	38	44	8.9	8.4	2.2	2.2	1.5	2.0	2.4	2.5	64	64
G-SEAT	Mean	165	156	-3.6	-1.8	9.5	-8.8	7.4	7.7	-1.6	-1.6		
	Disp	43	41	9.6	7.5	2.0	2.4	1.6	1.9	2.6	2.4	64	64
TURBULENCE	Mean	150	171*	-1.0	-4.3*	8.7	9.5*	8.2	6.9*	-1.1	-2.1*		
	Disp	44	37	8.2	8.8	1.9	2.4	1.7	1.7	2.2	2.6	64	64

[1] Values shown are for distance from the ramp in feet. The # 3 wire is at 186 ft. and the distance between the wires is approximately 33 feet.
[2] Refers to the number of trials that terminated in a touchdown on the landing deck.
* $p \leq .05$ for the mean difference between levels

G-Seat. Whenever G-seat had a statistically significant effect, performance was better with the G-seat off. During the descent phase, the percent time on target for glideslope during the last 1000 feet to the ramp was 9% larger with the G-seat off than with it on (Table 15). Lineup RMS error between 4500 and 2000 feet was also smaller with the G-seat off (Table 16). During touchdown, the lineup quality score was 10 units higher (better) when the G-seat was off (Table 18). When the G-seat was turned off, 88% of the 64 landings were successful; when it was on, 80% of 64 landings were successful. This difference is not statistically signficant ($p = .23$).

Ship Type. Wherever Ship Type showed a statistically significant effect, performance with the model-board ship was poorer than that with the computer generated ship. However, this occurred only with lineup performance during the descent (Table 16). Between 4500 and 500 feet from the ramp, the RMS lineup error was higher with the model-board ship than with the computer generated ship image. With the CIG ship type, 80% of the landings were successful. With the model-board ship, 88% of the landings were successful. This difference was not statistically significant.

Ship Type by G-Seat Interaction. During the descent (4500' - 2000'), a ship type by G-seat interaction occurred with RMS lineup error, signficant at $p = .01$ and accounting for 3.1% of the total variance (Table 16). For three of the four combinations, the RMS error was approximately 26°. For the fourth, when the G-seat was operating and the model-board carrier was used, the RMS lineup error increased to 40°. No such interaction effect was detected for this interaction when successful landings were considered ($p = .81$).

Other Factors. During the descent phase, turbulence showed little effect on lineup RMS error (Tables 16) but a significant effect on glideslope and angle of attack with RMS error lower with turbulence off (Tables 15 and 17).

When the percent of successful landings was considered, neither turbulence and its interactions with ship type and G-seat nor pilots showed any statistically significant effects. However, when only landing quality in the longitudinal direction (wire- trapped touchdown quality score, Table 18) was isolated, turbulence showed a large and surprising effect. The wire-trapped quality score was 72 with the turbulence off and 87 with the turbulence on. This difference was statistically significant ($p = .005$).

To examine this apparent reversal, we must consider the relatively large and statistically significant G-seat by turbulence interaction for the wire-trapped quality score (Table 18). The mean scores for the four combinations are:

		G-Seat OFF	G-Seat ON
Turbulence	ON	92	82
	OFF	65	80

Since other measures in Experiment III would lead us to expect the off-off cell to have the highest score, this result is surprising. It is seen that the observed interaction for the wire-trapped quality score is primarily the result of the off-off cell having the lowest quality score. Without further investigation, no interpretation of this result will be attempted.

SECTION VI

ANALYSES OF PILOT DATA

PILOT OPINIONS

Pilots' opinions regarding fidelity, adequacy for training, and adequacy for skill retention of the various equipment configurations employed in these experiments are summarized in Table 20. Mean ratings are shown for each level of each equipment factor; the higher the value, on a scale from 1 to 7, the better the rating. The median rating values are shown at the bottom of the table for each level across all factors (see Table 1 for the actual conditions).

While subjective evaluations of this type must be interpreted cautiously, some generalizations can be drawn from the table. For every factor, the mean rating among pilots was higher for the high level configuration than for the low level. This means that on the average they perceived the high level as being more realistic, better for training, and better for skill retention than the low level. The difference in many cases is slight. In some cases, the pilot preference--while agreeing with the experimenters' a priori judgments --does not necessarily agree with the empirical data, and frequently contrary opinions were expressed among the pilots. The data do not justify applying inferential statistics.

PILOT EXPERIENCE VERSUS SIMULATOR PERFORMANCE

In Table 21, the total number of actual flight hours, total number of actual carrier landings, and percent successful landings in the simulator are shown. When correlations were made between these parameters, the relationships obtained are remarkably strong considering the sample size. Rank correlations (Spearman's rho) are shown in Table 22.

TABLE 20. SUMMARY OF PILOT OPINIONS

Mean Ratings[1]

FACTOR	Fidelity		Adequacy For Training		Adequacy For Skill Retention	
Level[2]	Low	High	Low	High	Low	High
FLOLS	4.6	5.1	5.0	5.4	5.5	5.8
Ship Detail	3.6	5.4	3.5	5.6	4.4	5.6
Motion	3.4	5.1	3.9	5.4	4.5	5.5
Line Rate	3.6	5.6	3.8	5.8	4.0	5.8
Seascape	3.6	4.9	4.1	5.4	4.5	5.3
Brightness	4.1	5.3	4.3	5.4	4.8	5.4
Engine Lags	4.0	5.1	3.9	4.7	3.7	5.0
Ship Type	4.9	5.8	5.1	6.0	5.4	5.8
Visual Lags	3.9	4.9	4.1	5.1	4.4	5.0
FOV	4.1	5.4	3.9	5.6	4.7	5.7
G-Seat	2.6	3.9	2.8	3.9	2.8	3.9
Median	3.6	5.1	3.9	5.4	4.4	5.4

[1] Each level of each factor rated on scale of 1 - 7:

Very Poor					Very Good	
1	2	3	4	5	6	7

[2] Refer to Table 1 for description of high and low levels of factors.

TABLE 21. PILOT EXPERIENCE AND SIMULATOR PERFORMANCE

Pilot Experience		Percent Successful Simulator Landings			
Flight Hours	Carrier Landings	Exp. I	Exp. II	Exp. III	Combined I, II & III
4500	800	88	84	88	87
3700	400	38	59	81	59
2900	410	59	59	94	71
2500	400	81	63	94	79
2100	420	81	72	100	84
850	150	38	56	69	54
850	130	50	56	69	58
630	50	41	53	75	56

TABLE 22. RANK CORRELATIONS BETWEEN PILOT EXPERIENCE VARIABLES AND PERCENT SUCCESSFUL SIMULATOR LANDINGS (N = 8)

Pair being Correlated	Rank Correlations	Significance Levels (<p)
No. Flight Hours vs. Percent Successful Landings in Simulator:		
Experiment I	.62	.15
Experiment II	.78	.05
Experiment III	.49	.15
Combined I, II & III	.72 (.62)*	.10 (.10)
No. Carrier Landings vs. Percent Successful Landings in Simulator:		
Experiment I	.84	.02
Experiment II	.94	.01
Experiment III	.75	.05
Combined I, II & III	.89 (.82)*	.01 (.02)
Percent Successful Landings:		
Experiment I vs. Experiment II	.94	.01
Experiment I vs. Experiment III	.79	.05
Experiment II vs. Experiment III	.77	.05

* Pearson product-moment correlations in parenthesis.

SECTION VII

SUMMARY AND CONCLUSIONS

Studies were conducted to define design requirements for pilot training simulators for skill maintenance and transition training. Eleven pairs of simulator components were experimentally compared in the simulator to determine their effect on pilots with fleet experience in carrier landings. This study does not pertain to the training of novice pilots.

The main conclusion to be drawn from the results from the three experiments is that, on the whole, differences between the two levels of the 11 simulator factors investigated must be judged relatively smal' from a practical point of view. Some differences were statistically significant and probably reliable, but under the circumstances of these studies, individual differences (pilot experience) had a larger effect on performance than equipment configurations selected from among reasonable alternatives.

These conclusions are not particularly unexpected in view of the context in which these studies were performed. Factor levels, being selected from "reasonable alternatives"—one near the level presently existing in many carrier-landing trainers and the other representing a more advanced state of the art--are within operable ranges. The pilots themselves perceived only minor differences between the equipment levels insofar as fidelity, adequacy of training, and maintaining skills are concerned. Then too, pilots who participated in the study had considerable flight experience as well as actual carrier-landing experience. As a rule, skilled pilots are highly adaptable, and would be expected to perform adequately under less than optimum conditions.

In short, despite the fact that (a) the pilots were experienced, and (b) the equipment factors were examined over a wide range of practical interest, it is clear that performance differences among pilots greatly overshadowed the differences due to equipment. This implies that a point of diminishing returns has been reached with respect to further improvements in simulator fidelity for carrier landing skill maintenance and transition training. It means that in future trainers for this task, added realism should not be purchased at the expense of lower reliability or higher acquisition and life-cycle costs.

INDIVIDUAL FACTORS

In spite of this basic conclusion, it is interesting to note that all of the factors had some effect on something. Certain conditions and configurations among the eleven investigated affected pilot performance in a way that definitely enhanced the mission of landing safely on the carrier. Some factors had only marginal effects or appeared to affect only secondary criteria in the landing process, while others showed statistical but not necessarily practical effects. Some results were contrary to what had been expected. A brief summary of results follows with factor effects listed in the order of estimated overall impact on the simulated carrier landing task.

PRIMARY EFFECTS. The factors showing the largest performance differences between conditions were FLOLS and ship detail.

The CIG FLOLS resulted in a significantly smaller vertical error along the glideslope than the Optical FLOLS during the descent to the carrier. There was also evidence that there was less longitudinal dispersion at touchdown with the CIG FLOLS. Since performance with both devices fell within acceptable limits, there was no significant difference in the number of successful landings made with each. No comparison of the two devices was made during the turn approach.

Ship detail had a significant effect on lineup with the high-detail ship producing better performance. With the low-detail ship, pilots tended to get consistently right of centerline in the middle of the approach, often resulting in an overcorrection near the ramp and an earlier, slightly off-center touchdown. With the circular approach, the same pattern was noted although it was not as pronounced. In spite of the observed differences, approaches with both configurations were within acceptable limits. Still, overall, the high-detail ship resulted in more successful landings than did the low-detail ship (3% more, straight-in; 14% more, circling).

MARGINAL EFFECTS. Marginal or secondary effects were observed between levels of visual lag, line rate, field of view, seascape, and engine lag. Secondary effects are those that reliably affected the position of the aircraft about its own axis but not the aircraft's position along the optimum path of descent and at touchdown. These effects involve the pilots' ability to maintain a stable aircraft during flight, but not to the extent that the outcome of the task is significantly affected. Marginal effects are those that were moderate in size but whose reliability was questionable, depending on the criteria with which the data were interpreted.

The shorter (100 msec) visual lag resulted in significantly less roll variability during the approach than the 200 msec visual lag. The shorter lag also resulted in more successful landings (3% more, straight-in; 12% more, circling). The difference in successful landings for circling approaches was marginally significant.

Line rate showed an effect contrary to what was expected a priori: 14% more successful landings were made with the 525 line system than with the 1025. The difference was marginally significant. The data revealed that with the 1025 line rate, there was an average tendency in close to the ramp to overcorrect for right biased lineup errors, with the result that fewer touchdowns were made within acceptable tolerance limits. The effect was marginal and not supported by other measures of performance. No comparison of the two line rates was made during the circular approach.

The presence of seascape produced a bias in judging distance from the ship during the turn of the circling approach, a lower angle of attack during the final approach and harder landings at touchdown. As a result, 10% fewer successful touchdowns were made with seascape on than when it was off. No such extensive biases were observed during the straight-in approach, and there

were 11% more successful landings with the seascape on with straight-in approaches. Apparently, the seascape produced a bias error in the turn which never materialized with straight-in approaches. The differences in successful landings were not statistically significant.

The wide field of view resulted in significantly less roll variability during descent and at touchdown than the narrow field of view. There was essentially no difference in the number of successful landings made during the straight-in approaches with the two conditions. No comparison of fields of view were made during the circling approach.

With the faster engine update rate of 30 Hz (i.e., shorter engine lag), pitch was significantly less variable close to the ramp. The faster update also resulted in a slightly higher percentage of successful landings, but the three percent difference was not even marginally significant. No comparison of engine lags was made with circling approaches.

SMALLER EFFECTS. The remaining effects, while statistically detectable for certain measures, are probably too small to be of practical value; or, because of inconsistencies and lack of supporting evidence must be viewed with considerable caution.

Glideslope control close to the ramp was slightly better with the G-seat off. This result is contrary to the expectation that the G-seat would provide more cues and facilitate performance. Also, there were eight percent more successful landings with the G-seat off than with it on, although the difference is not statistically significant. No comparison was made with the circling approaches.

The CIG ship type resulted in slightly better lineup performance during the descent phase but there were eight percent more successful landings with the model-board ship type, although this latter difference was not significant. No comparison was made for circling approaches.

High brightness resulted in less roll variability during the descent phase and in more successful landings (8% straight-in; 0% circling). These latter differences were not statistically significant.

Virtually no practical effect of motion was observed except for some minor biases. The differences in number of successful landings between motion on and off was small and inconsistent between experiments.

CONSIDERATIONS

By citing marginal results, more attention may be drawn to them than is justified. When an observed difference is not statistically significant in these experiments, assuming the probability level chosen for significance is reasonable, then it is quite likely that it is not reliable, for care was taken to include enough observations to provide the power required to draw that conclusion. Each mean difference between the two levels of every factor

was based on 128 pairs of observations in the first two experiments and 64 pairs in the third.

Still one must not discard factors too quickly. Small differences may expand (or disappear) when the difficulty level shifts, as when less experienced pilots or more difficult tasks are employed. Also, small reliable differences when summed over several factors may become large enough to have operational significance. For example, if five factors each accounted for only 3% of the total variance, that might not be too important individually, but would be when considered together, accounting for 15% of the total variance.

Finally, it should be noted that no single simulation study is sufficient by itself. While the first two experiments gave more information than could have been obtained with at least 45 two-factor experiments with 256 observations each, they only answer the specific question regarding performance differences between configurations in the simulator. A simulator-to-field transfer-of-training experiment is ultimately needed.

REFERENCES

Box, G.E.P. and Hunter, J.S. The 2^{k-p} fractional factorial designs. Part I. Technometrics, 1961, 3, 311-351.

Brictson, C.A. Burger, W.J., and Wulfeck, J.W. Validation and application of a carrier landing performance score: the LPS. Arlington, VA: Office of Naval Research, TR NR 196-115, Contract No. N00014-72-C-0041, 1973.

Browder, G.B. and Butrimas, S.K. Visual Technology Research Simulator Visual and Motion System Dynamics. Orlando, FL: Naval Training Equipment Center TR NAVTRAEQUIPCEN IH-326, April 1981.

Collyer, S.C. and Chambers, W.S. AWAVS, A Research Facility for Defining Flight Trainer Visual System Requirements. Proceeding of the Human Factors Society 22nd Annual Meeting, Detroit, Michigan: Human Factors Society, 1978.

Davies, O.L. (Ed) The Design and Analysis of Industrial Experiments. Hafner Publishing Co., New York, 1967.

General Electric Co. System description--Aviation wide-angle visual system (AWAVS) computer image generator (CIG) visual system, Space Division, Daytona Beach, FL: TR NAVTRAEQUIPCEN 76-C-0048-1, February 1979.

Golovcsenko, I.V. Computer simulation of Fresnel Lens Optical Landing System, Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN IH-265, September 1976.

Kaul, C.E., Collyer, S.C. and Lintern, G. Glideslope descent-rate cuing to aid carrier landings. Orlando, FL: Naval Training Equipment Center, NAVTRAEQUIPCEN IH-322, October 1980.

Lintern, G., Westra, D., Iavecchia, H., and Hennessy, R. Visual Technology Research Simulator (VTRS) Human Performance Research: Phase II. Canyon Research Group, Inc., Westlake Village, CA: TR NAVTRAEQUIPCEN 78-C-0060-2, April, 1980.

Lintern, G., Westra, D., Iavecchia, H. and Roscoe, S.N. Visual Technology Research Simulator (VTRS) Human Performance Research: Phase I. Canyon Research Group, Inc., Westlake Village, CA: TR NAVTREQUIPCEN 78-C-0060-1, April 1979.

National Bureau of Standards. Fractional factorial experiment designs for factors at two levels. Stat Eng. Lab., Applied Math Series No. 48, U.S. Government Printing Office, Washington, DC: 15 April 1957.

Orlansky, J. and String, J. Cost-effectiveness of flight simulators for military training. Volume I: Use and effectiveness of flight simulators. Institute for Defense Analyses, Science and Technology Division, IDA Paper P-1275, 1977.

Owen, J.R. Distribution of monochrome screen luminance in the CTOL Visual Technology Research Simulator. Orlando, FL: Naval Training Equipment Center TR NAVTRAEQUIPCEN IH-328, November 1980.

Simon, C.W. Economical multifactor designs for human factors engineering experiments. Culver City, CA: Hughes Aircraft Co., TR P73-326A, June 1973 (AD A035-108).

Simon, C.W. New research paradigm for applied experimental psychology: a system approach. Westlake Village, CA: Canyon Research Group, Inc., TR CWS-04-77A, October, 1977 [Revised June 1978] (AD A056-984).

Singer-Link Division. Aviation wide-angle visual trainer subsystem design report, Binghampton, NY: TR NAVTRAEQUIPCEN 75-C-0009-13, May 1977.

Westra, D.P. Investigation of simulator design features for the carrier-landing task: An in-simulator transfer of training experiment. Canyon Research Group, Inc., West lake Village, CA: NAVTRAEQUIPCEN 81-C-0105-1, In Press.

APPENDIX A

TABLE A1. LIST OF PERFORMANCE MEASURES: WITHIN TRIAL SUMMARY SCORES

Final Approach Aircraft Output and Outcome Summaries

Task Segments	Variables Measured in Each Segment	Summary Measures of Each Variable
FLOLS Space Entry to Rollout/ 6000–4500 Feet[1]	Angle-of-Attack Deviation (Units)	RMS = $(\Sigma\, e_t^2 / n)^{1/2}$
Rollout – 3000 Feet/4500–2000 Feet[1]	Roll (Degrees)	Bias = $\bar{e}$
3000–1000 Feet	Pitch (Degrees)	Variability = $(\Sigma\, (e - e_t^2 / n)^{1/2}$
2000–500 Feet	Lineup Deviation (feet and degrees)	Percent Time on Target:
1000 Feet–Ramp	Glidescope Deviation (feet and degrees)	Angle-of-Attack = ± 1.0 Unit Lineup = ± 1.5 Degrees Glidescope = ± .3 Degree

Stick Movement Scores

Task Segments	Variables Measured in Each Segment	Summary Measure of Each Variable
FLOLS Space Entry to Rollout/ Start to 4500 Feet[1]	Throttle	Average Stick Movement per second = $r/n\ \Sigma(P_t - P_{t-1})$ where P_t = stick position at time t, r = sampling rate = 30 Hz, n = total no. samples
Rollout – 1600 Feet/4500–1600 Feet[1]	Aileron	
1600–400 Feet	Elevator	
400 Feet to Touchdown	Pedal	

[1]First segment applies to circling approaches, second to straight-in approaches.

APPENDIX A

TABLE A2. LIST OF PERFORMANCE MEASURES: APPROACH CAPTURE VARIABLES

Final Turn Captures	Variables Measured at Each Capture Point
0° (Start of final turn)	Distance From Ship Centerline (Feet)
22.5°	Distance Aft Ship Center of Gravity (Feet)
45°	Aircraft Altitude (Feet)
67.5°	Angle-of-Attack (Units)
90° (Aft Ship on Centerline)	Bank Angle (Degrees)
Final Approach Captures	**Variables Measured at Each Capture Point**
FLOLS Space Entry/Start	Angle-of-Attack Deviation (Units)
Rollout/4500 Feet	Lineup Deviation (Feet)
3000 Feet	Glideslope Deviation (Feet)
2000 Feet	Roll Angle (Degrees)
1000 Feet	Pitch Angle (Feet)

First segment refers to circling approaches, second to straight-in approaches.

TABLE A3. LIST OF PERFORMANCE MEASURES:

TOUCHDOWN CAPTURE VARIABLES

Touchdown Capture Variables

Wire Trapped

Distance From Ramp (Feet)

Distance From Center of Landing Deck (Feet)

Aircraft Roll (Degrees)

Aircraft Pitch (Degrees)

Vertical Velocity (Ft/Sec)

Angle of Attack (Units)